Linear Algebra II

陳信睿

January, 2023

Abstract

這篇筆記主要是因爲在預習線性代數二的時候，常常發現很多重要的定理都記不太起來，並且老師在下學期沒有選定指定的參考書，所以我就寫了這份筆記。主要是參考謝銘倫老師的影片 [2]，以及著名的線性代數教科書 [1] 所寫。

內容目前涵蓋了商空間、對偶空間以及內積空間的大部分內容，甚至比 “Linear Algebra”[1] 中還要多東西，像是 Hilbert space。不過我盡量把證明寫的精簡一點，同時我也省去了所有的範例。

I wrote this note because I often found that I could not remember many important theorems when I was studying Linear Algebra II, and my teacher did not choose a reference book for the next semester. The main reference is Professor Ming-Lun Hsieh’s video [2], and the famous linear algebra textbook [1].

The content now covers most of the quotient space, dual space, and inner product space, even more than in “Linear Algebra”[1], like Hilbert space. I have tried to keep the proof as concise as possible, and I have also omitted all the examples.

Contents

1 Quotient and dual spaces

1.1 Quotient space

Definition 1 (Quotient space)**.** Let V be a vector space and let W be its subspace. Define an equivalence relation on V such that

$$v_1 \sim v_2 \text{ if } v_1 - v_2 \in W.$$

It is easy to verify that $\sim$ is indeed an equivalence relationship on V. For each $v_0 \in V$, define $[v_0] = \{v \in V : v \sim v_0\}$ the equivalence class of v_0. Then, $\{[v] : v \in V\}$ is called the quotient space V/W.

Remark. The quotient space V/W is equipped with a natural vector (linear) structure, namely,

$$\begin{cases} [v_1] + [v_2] = [v_1 + v_2] & \text{, for all } v_1, v_2 \in V \\ c[v_1] = [cv_1] & \text{, for all } v_1 \in V \text{ and } c \in \mathbb{F} \end{cases} .$$

Although it is crucial that we shall check these natural addition and scalar multiplication are "well-defined", we omitted here.

Definition 2 (Quotient maps)**.** There is a natural surjective map

$$\begin{aligned} \pi : V &\to V/W \\ v &\mapsto [v] \end{aligned} ,$$

which is called the quotient map. Moreover, it is a linear transformation.

Remark.

$$\begin{aligned} \ker \pi &= \{v \in V : \pi(v) = [0]\} \\ &= \{v \in V : [v] = [0]\} \\ &= \{v \in V : v - 0 \in W\} \\ &= W \ . \end{aligned}$$

Corollary. It follows from the dimension formula that $\dim_{\mathbb{F}} V/W = \dim_{\mathbb{F}} V - \dim_{\mathbb{F}} W$ whenever V is finite dimensional.

Here we give an alternative proof without using dimensional formula. Since V has finite dimension, let $\mathcal{B} = \{w_1, w_2, \ldots, w_s\}$ be a basis of W and extend $\mathcal{B}$ to $\mathcal{A} = \{w_1, w_2, \ldots, w_r\}$ a basis of V. We claim that $\{[w_{s+1}], \ldots, [w_s]\}$ is a basis of V/W. To see this, we shall show that:

1. $\{[w_{s+1}], \ldots, [w_r]\}$ generate V/W.
   Suppose $[v] \in V/W$. Let $v = \sum_{i=1}^{r} \alpha_i w_i$, then

$$[v] = \left[\sum_{i=s+1}^{r} \alpha_i w_i \right] = \sum_{i=s+1}^{r} \alpha_i [w_i] \quad .$$

2. $\{[w_{s+1}], \ldots, [w_r]\}$ are linear independent over $\mathbb{F}$.
Suppose $\sum_{i=s+1}^{r} \alpha_i \cdot [w_i] = [0]$, for some $\alpha_i \in \mathbb{F}$. Then,

$$\begin{aligned} & \left[\sum_{i=s+1}^{r} \alpha_i w_i\right] = [0] \\ \iff & \sum_{i=s+1}^{r} \alpha_i w_i \in W \\ \iff & \sum_{i=s+1}^{r} \alpha_i w_i = \sum_{j=1}^{s} \beta_j w_j, \text{ for some } \beta_j \in \mathbb{F}. \end{aligned}$$

We conclude that α_i are all zeros, since $\mathcal{A}$ is a basis of V.

Discussions above show that $\dim_{\mathbb{F}} V/W = r - s = \dim_{\mathbb{F}} V - \dim_{\mathbb{F}} W$. Now, we shall study some property about the quotient space V/W. The next theorem characterize the quotient space V/W by the following universal property.

Theorem 3. *Let T be a linear transformation from V to U, such that* $\ker T$ *contain W, namely $W \subset \ker T$. Then, T factors through π uniquely. That is, there exists a unique linear transformation $S : V/W \to U$ such that*

$$T = S \circ \pi.$$

Proof. Define $S : V/W \to U$ by

$$S([v]) = T(v).$$

We first show that S is a well-defined map, namely, if $[v] = [v']$, then $T(v) = T(v')$. Note that $[v] = [v'] \implies v - v' \in W \subset \ker T$, we conclude $T(v) = T(v')$. By definition, S is a linear transformation and $S \circ \pi = T$. The uniqueness of such S follows from the surjectivity of π. □

Remark. The quotient space V/W with the quotient map π is the unique vector space satisfying the theorem. That is, if we are given $\pi' : V \to V'$ satisfying the property: for every linear transformation $T : V \to U$ with $W \subset \ker T$, there exists a unique $S' : V' \to U$ such that $S' \circ \pi' = T$. Then, $V' \simeq V/W$ uniquely.

Proof. From the assumptions, we have

$$\begin{cases} \exists!\ S : V/W \to V', \text{ such that } \pi' = S \circ \pi \\ \exists!\ S' : V' \to V/W, \text{ such that } \pi = S' \circ \pi' \end{cases}.$$

This shows $S \circ S' = \mathrm{Id}_{V'}$; $S' \circ S = \mathrm{Id}_{V/W}$ (using Theorem 3 again.) We conclude $V' \simeq V/W$ uniquely. □

Corollary. Let $T : V \to W$ be a linear transformation. Then,

$$V/\ker T \simeq \mathrm{Im} T.$$

Hence, $\dim_{\mathbb{F}} V/\ker T = \dim_{\mathbb{F}} \mathrm{Im} T$.

Proof. From Theorem 3, we have: there exists a unique $S : V/\ker T \to W$, such that $T = S \circ \pi$. It follows from the surjectivity of π that $\mathrm{Im}S = \mathrm{Im}T$. We claim that S is injective. Note that

$$\begin{aligned}\ker S &= \{[v] \in V/\ker T : S([v]) = 0\} \\ &= \{[v] \in V/\ker T : T(v) = 0\} \\ &= \{[v] \in V/\ker T : v \in \ker T\} \\ &= \{[0]\}\,.\end{aligned}$$

Thus, S is a bijection. This completes the proof. □

Now, let $T : V \to V$ be a linear transformation and let $W \subset V$ be a T-invariant subspace. Then, T induce a linear transformation $\widetilde{T}$ on V/W define by:

$$\begin{aligned}\widetilde{T} : V/W &\to V/W \\ [v] &\mapsto [T(v)]\end{aligned}\,.$$

This is a well-defined map since

$$\begin{aligned}[v] = [v'] &\implies v - v' \in W \\ &\implies T(v) - T(v') = T(v - v') \in W \\ &\implies [T(v)] = [T(v')].\end{aligned}$$

Now, let $\mathcal{B} = \{v_1, v_2, \ldots, v_s\}$ be a basis of W, and extend it to $\mathcal{A} = \mathcal{B} \sqcup \mathcal{B}'$, a basis of V. We have shown that $[\mathcal{B}'] = \{[v] : v \in \mathcal{B}'\}$ is a basis of V/W. Then, we have

$$[T]_{\mathcal{A}} = \left(\begin{array}{c|c} [T|_W]_{\mathcal{B}} & * \\ \hline 0 & \left[\widetilde{T}\right]_{[\mathcal{B}']} \end{array}\right).$$

We thus have

$$\begin{cases}\mathrm{ch}_T(x) = \mathrm{ch}_{T|_W}(x) \cdot \mathrm{ch}_{\widetilde{T}}(x) \\ \mathrm{m}_T(x) \text{ is divisible by } \mathrm{m}_{T|_W}(x)\end{cases}.$$

Corollary. If T is diagonalizable, then so is $\widetilde{T}$.

The corollary follows from the fact that $\mathrm{m}_T(x)$ is divisible by $\mathrm{m}_{\widetilde{T}}(x)$. We next shall discuss the concept of dual spaces.

1.2 Dual space

Definition 4 (dual space)**.** Let V be a vector space over $\mathbb{F}$. It is well-known that $L(V, \mathbb{F})$ is a vector space over $\mathbb{F}$. It is called the dual space of V, and its elements are called linear functionals of V. We often write $V^\vee$ to denote the dual space of V.

Recall that:

Given two vector spaces V, W over $\mathbb{F}$. Then we have $L(V, W)$ is a vector space over $\mathbb{F}$ and

$$\dim_{\mathbb{F}} L(V, W) = \dim_{\mathbb{F}} V \cdot \dim_{\mathbb{F}} W.$$

Thus, we conclude that $\dim_{\mathbb{F}} V^{\vee} = \dim_{\mathbb{F}} V$ if $\dim_{\mathbb{F}} V < \infty$. Here we give an alternative proof.

Theorem 5. *Suppose V is a finite dimensional vector space over $\mathbb{F}$. Then, $\dim_{\mathbb{F}} V^{\vee} = \dim_{\mathbb{F}} V$.*

Proof. Let $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ be a basis of V. Let us consider the following linear functional:

$$\begin{aligned} v_i^{\vee} : \ V &\to \mathbb{F} \\ \sum_{i=1}^{n} \alpha_i \cdot v_i &\mapsto \alpha_i \end{aligned}$$

We claim that $\mathcal{B}^{\vee} = \{v_1^{\vee}, v_2^{\vee}, \ldots, v_n^{\vee}\}$ is a basis of $V^{\vee}$, the dual space of V. We first show that $\mathcal{B}^{\vee}$ is linear independent. Suppose there exist $\beta_i \in \mathbb{F}$ such that

$$\sum_{i=1}^{n} \beta_i v_i^{\vee} = 0,$$

then

$$\sum_{i=1}^{n} \beta_i v_i^{\vee}(v_j) = 0.$$

This shows

$$\beta_i = 0, \text{ for all } i = 1, 2, \ldots, n.$$

Next we show that $\mathcal{B}^{\vee}$ generate $V^{\vee}$. Given $l \in V^{\vee}$. Then, from the linearity of l, we have

$$l = \sum_{i=1}^{n} l(v_i) \cdot v_i^{\vee}.$$

We conclude that $\mathcal{B}^{\vee}$ is a basis of $V^{\vee}$. $\square$

Remark. The basis $\mathcal{B}^{\vee}$ is called the dual basis of $\mathcal{B}$.

Given a linear transformation $T : V \to W$, it induces a linear transformation $T^{\vee} : W^{\vee} \to V^{\vee}$ between dual spaces defined by:

$$T^{\vee}(l)(v) := l(T(v)), \text{ for } l \in W^{\vee} \text{ and } v \in V.$$

It is easy to verify that $T^{\vee}$ is a linear transformation.

Theorem 6. *Let V, W be two finite dimensional vector spaces over $\mathbb{F}$. Let $\mathcal{A} = \{v_1, v_2, \ldots, v_n\}$ and $\mathcal{B} = \{w_1, w_2, \ldots, w_m\}$ be bases of V and W, respectively. Given $T : V \to W$. Then,*

$$[T]^{\mathrm{t}}_{\mathcal{A},\mathcal{B}} = [T^{\vee}]_{\mathcal{B}^{\vee},\mathcal{A}^{\vee}}.$$

Proof. Let $A := [T]_{\mathcal{A},\mathcal{B}} = (a_{ij})_{n\times n}$ and $B := [T^{\vee}]_{\mathcal{B}^{\vee},\mathcal{A}^{\vee}} = (b_{ij})_{n\times n}$. From the definition, we have

$$\begin{aligned} T(v_j) &= \sum_{i=1}^{m} a_{ij} w_i \\ T^{\vee}(w_i^{\vee}) &= \sum_{j=1}^{n} b_{ji} v_j^{\vee} \end{aligned}.$$

Then,

$$b_{ji} = T^\vee(w_i^\vee)(v_j) = w_i^\vee(T(v_j)) = w_i^\vee\left(\sum_{i=1}^{m} a_{ij}w_i\right) = a_{ij}.$$

This proves the theorem. $\square$

Theorem 7. *Let V be a vector space and let $W \subset V$ be a subspace. Then,*

$$(V/W)^\vee \simeq \{l \in V^\vee : W \subset \ker l\}.$$

Proof. We have known that there is a natural map $\pi : V \twoheadrightarrow V/W$. We claim that $\pi^\vee$ is the isomorphism that bijects $(V/W)^\vee$ and $\{l \in V^\vee : W \subset \ker l\}$. We first show that $\pi^\vee$ is injective. Suppose $\pi^\vee(l) = 0$, for some $l \in (V/W)^\vee$. Then,

$$\begin{aligned} &l(\pi(v)) = 0, \text{ for all } v \in V \\ \Longrightarrow\ &l([v]) = 0, \text{ for all } v \in V. \end{aligned}$$

This shows the injectivity of $\pi^\vee$. Hence, $(V/W)^\vee \simeq \mathrm{Im}\pi^\vee$. It suffices to show that $\mathrm{Im}\pi^\vee = \{l \in V^\vee : W \subset \ker l\}$.

1. $\mathrm{Im}\pi^\vee \subset \{l \in V^\vee : W \subset \ker l\}$.
   For each $S \in (V/W)^\vee$ and $w \in W$, we have
   $$\pi^\vee(S)(w) = S(\pi(w)) = S([w]) = S([0]) = 0.$$
2. $\{l \in V^\vee : W \subset \ker l\} \subset \mathrm{Im}\pi^\vee$.
   Let $l \in V^\vee$ such that $W \subset \ker l$. Theorem 3 asserts that there exists a unique $S : V/W \to \mathbb{F}$ such that $l = S \circ \pi$. This implies $\pi^\vee(S) = l$.

Discussions above complete the proof. $\square$

Corollary. Let $A \in M_{m\times n}(\mathbb{F})$. Then, $\mathrm{rank}A = \mathrm{rank}A^{\mathrm{t}}$.

Proof. Let $V = \mathbb{F}^n$, $W = \mathbb{F}^m$ and let $T : V \to W$ defined by

$$T(v) = A \cdot v.$$

Then it is equivalent to prove

$$\dim \mathrm{Im}T = \dim(\mathrm{Im}T^\vee).$$

By Theorem 7,

$$(W/\mathrm{Im}T)^\vee \simeq \{l \in W^\vee : \mathrm{Im}T \subset \ker l\} = \{l \in W^\vee : T^\vee(l) = 0\} = \ker(T^\vee). \tag{1}$$

Thus,

$$\dim W - \dim \mathrm{Im}T = \dim W/\mathrm{Im}T = \dim(W/\mathrm{Im}T)^\vee = \dim W^\vee - \dim \mathrm{Im}(T^\vee).$$

This completes the proof. $\square$

Theorem 8. *Let V and W are two finite vector spaces, and let $T : V \to W$ be a linear transformation. Then,*

1. *T is surjective if and only if $T^\vee$ is injective.*

2. *T is injective if and only if $T^\vee$ is surjective.*

Proof. In the proof of the previous corollary, we have shown in equation 1 that

$$(W/\mathrm{Im}T)^\vee \simeq \ker(T^\vee),$$

this proves the first assertion. Similarly, we have

$$(V/\ker T)^\vee \simeq \{l \in V^\vee : \ker T \subset \ker l\}. \tag{2}$$

We claim the set on the right hand side is $\mathrm{Im}(T^\vee)$.

1. $\{l \in V^\vee : \ker T \subset \ker l\} \subset \mathrm{Im}(T^\vee)$.
   Let $l \in V^\vee$ such that $\ker T \subset \ker l$. It is well-known that there exist a subspace $X \subset W$ such that $W = \mathrm{Im}T \oplus X$. Consider a transformation $s : W \to \mathbb{F}$ defined by:
   $$s(w) = l(v),$$
   where $w = T(v) + x$, for some $v \in V$ and $x \in X$. This is a well-defined map, since $\ker T \subset \ker l$. Note that s is a linear transformation and $l = s \circ T = T^\vee(s)$. This implies $\{l \in V^\vee : \ker T \subset \ker l\} \subset \mathrm{Im}(T^\vee)$.

2. $\mathrm{Im}(T^\vee) \subset \{l \in V^\vee : \ker T \subset \ker l\}$.
   Let $l =\in \mathrm{Im}(T^\vee)$. Then, there exists $s \in W^\vee$ such that $l = T^\vee(s) = s \circ T$, thus $\ker T \subset \ker l$.

Discussions above with equation 2 show that

$$(V/\ker T)^\vee \simeq \mathrm{Im}(T^\vee),$$

which is equivalent to the second assertion. □

Remark. In the class, the teacher prove with another approach, which use the following property:

> Let V be a finite dimensional vector space, and let $V^{\vee\vee}$ be the dual space of V, then there is a natural identification, that is, there is an isomorphism $\phi : V \to V^{\vee\vee}$ defined by
>
> $$\phi : x \mapsto (\hat{x} : f \mapsto f(x)), \quad f \in V^\vee.$$

Next, we show that why we shall study dual spaces by the following theorem.

Theorem 9. *Let V be a finite dimensional vector space over $\mathbb{F}$. Let $l_1, l_2, \ldots, l_s \in V^\vee$ be linearly independent. Suppose $b_1, b_2, \ldots, b_s \in \mathbb{F}$ and put*

$$\Xi = \{v \in V : l_i(v) = b_i, \text{ for all } 1 \le i \le s\}.$$

Then, $\Xi \neq \varnothing$.

Proof. Consider the linear transformation $T : V \to \mathbb{F}^s$ defined by:

$$T : v \mapsto (l_1(v), l_2(v), \ldots, l_s(v)).$$

Then, $\dim \ker T$ is $\dim V - s$. Here we omit the details of the proof. □

2 Inner product space

Definition 10 (inner product). Let V be a vector space over $\mathbb{F}$, where $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ is called an inner product if the following conditions are satisfied:

1. $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$, for all $x, y, z \in V$.
2. $\langle cx, y \rangle = c \cdot \langle x, y \rangle$, for all $x, y \in V$ and $c \in \mathbb{F}$.
3. $\langle x, y \rangle = \overline{\langle y, x \rangle}$, for all $x, y \in V$.
4. $\langle x, x \rangle \geq 0$, for all $x \in V$ and $\langle x, x \rangle = 0$ if and only if $x = 0$.

We write $(V, \langle \cdot, \cdot \rangle)$ for a vector space V together with an inner product structure $\langle \cdot, \cdot \rangle$. In the following text, $\mathbb{F}$ sill stand for $\mathbb{R}$ or $\mathbb{C}$ unless otherwise stated.

We could also define the concept of norm or length of a vector $v \in V$.

Definition 11 (norm). For each $v \in V$, define the norm of v as $\|v\| = \langle v, v \rangle^{1/2}$.

Theorem 12 (Riesz representation Theorem on a finite dimensional space). *Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. Then,*

$$\begin{aligned} \Phi : V &\to V^\vee \\ v &\mapsto \Phi(v)(x) = \langle x, v \rangle \end{aligned}$$

is an isomorphism.

Proof. We first prove that Φ is injective. Note that

$$\ker \Phi = \{ v \in V : \langle x, v \rangle = 0, \text{ for all } x \in V \} = \{0\}.$$

Since V is finite dimensional, we have $\dim_{\mathbb{F}} V = \dim_{\mathbb{F}} V^\vee$, thus Φ is an isomorphism. □

In other words, inner product $\langle \cdot, \cdot \rangle$ identifies V with its dual space $V^\vee$ when V is finite dimensional. We now start study how to represent an inner product structure with a matrix. Suppose V is a finite dimensional vector space, and let $\mathcal{A} = \{v_1, v_2, \ldots, v_n\}$ be a basis of V. For any $x, y \in V$, there exist α_i, β_i such that

$$x = \sum_{i=1}^{n} \alpha_i \cdot v_i \,; \quad y = \sum_{j=1}^{n} \beta_j \cdot v_j.$$

Then,

$$\langle x, y \rangle = \left\langle \sum_{i=1}^{n} \alpha_i \cdot v_i, \sum_{j=1}^{n} \beta_j \cdot v_j \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \overline{\beta_j} \langle v_i, v_j \rangle.$$

Hence, if we let

$$\Omega = (\langle v_i, v_j \rangle) \in M_n(\mathbb{F}),$$

we have

$$\langle x, y \rangle = \begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix} \cdot \Omega \cdot \begin{pmatrix} \overline{\beta_1} \\ \overline{\beta_2} \\ \vdots \\ \overline{\beta_n} \end{pmatrix}.$$

The matrix Ω is called the matrix of $\langle \ , \ \rangle$ associated with $\mathcal{A}$.

Theorem 13 (change of basis). *Let $\mathcal{B} = \{w_1, \ldots, w_n\}$ be another basis of V. Assume that*

$$w_j = \sum_{i=1}^{n} a_{ij} v_i, \text{ for all } 1 \le j \le n.$$

Then,

$$\Omega' = A^{\mathrm{t}} \cdot \Omega \cdot \overline{A},$$

where Ω' is the matrix of $\langle\ ,\ \rangle$ associated with $\mathcal{B}$ and $A = (a_{ij})$.

Proof. Note that

$$\begin{aligned}
\langle w_i, w_j \rangle &= \left\langle \sum_{k=1}^{n} a_{ki} v_k, \sum_{l=1}^{n} a_{lj} v_l \right\rangle \\
&= \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ki} \langle v_k, v_l \rangle \overline{a_{lj}} \\
&= \sum_{k=1}^{n} \sum_{l=1}^{n} {a_{ik}}^{\mathrm{t}} \langle v_k, v_l \rangle \overline{a_{lj}},
\end{aligned}$$

This proves the theorem. $\square$

Next, we shall ask whether we can define an inner product structure on V if we are given a matrix $\Omega \in M_n(\mathbb{F})$ and a basis $\mathcal{A}$ of V. The answer is no. In fact, the matrix can define an inner product structure on finite dimensional V if and only if it is positive definite. However,

Theorem 14. *If $\Omega = B \cdot B^*$ for some $B \in M_n(\mathbb{F})$ with $\det B \neq 0$, then $\langle\ ,\ \rangle_{\Omega,\mathcal{A}}$ is an inner product for any choice of $\mathcal{A}$.*

Proof. Let $\mathcal{A} = \{v_1, v_2, \ldots, v_n\}$ be an arbitrary basis of V. It suffices to show the inner product defined by Ω satisfies the fourth axiom of Definition 10. If $x \in V$, then

$$x = \sum_{i=1}^{n} \alpha_i \cdot v_i, \text{ for some } \alpha_i \in \mathbb{F}.$$

We have

$$\begin{aligned}
\langle x, x \rangle_{\Omega,\mathcal{A}} &:= \begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix} \cdot \Omega \cdot \begin{pmatrix} \overline{\alpha_1} \\ \overline{\alpha_2} \\ \vdots \\ \overline{\alpha_n} \end{pmatrix} \\
&= \begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix} \cdot B \cdot B^* \cdot \begin{pmatrix} \overline{\alpha_1} \\ \overline{\alpha_2} \\ \vdots \\ \overline{\alpha_n} \end{pmatrix} \\
&= (yB) \cdot (yB)^*,
\end{aligned}$$

where $y = \begin{pmatrix} \alpha_1 & \alpha_2 & \ldots & \alpha_n \end{pmatrix}$ is a row vector. Write $yB = \begin{pmatrix} \beta_1 & \beta_2 & \ldots & \beta_n \end{pmatrix}$. We get

$$\langle x, x \rangle_{\Omega,\mathcal{A}} = \begin{pmatrix} \beta_1 & \beta_2 & \ldots & \beta_n \end{pmatrix} \cdot \begin{pmatrix} \overline{\beta_1} \\ \overline{\beta_2} \\ \vdots \\ \overline{\beta_n} \end{pmatrix} = \sum_{i=1}^{n} |\beta_i|^2 \ge 0,$$

and $\langle x, x\rangle_{\Omega,\mathcal{A}} = 0$ if and only if $y = 0$. From the assumption that $\det B \neq 0$, it follows $x = 0$ if $\langle x, x\rangle = 0$. $\square$

Definition 15 (Hermitian and positive definite matrix)**.** Let $\Omega \in M_n(\mathbb{F})$. Then,

1. Ω is said to be Hermitian if $\Omega^* = \Omega$.
2. Ω is said to be positive definite if Ω is Hermitian and

$$x \cdot \Omega \cdot x^* > 0, \text{ for all row vector } x \in \mathbb{F}^n \backslash \{0\}.$$

Remark. Let $\Omega \in M_n(\mathbb{F})$. Define an $\langle \cdot, \cdot \rangle$ on the vector space $V = \mathbb{F}^n$ by

$$\langle x, y\rangle = x \cdot \Omega \cdot y^*, \text{ where } x \text{ and } y \text{ are row vectors,}$$

then $\langle\ ,\ \rangle$ is an inner product on V if and only if Ω is positive definite.

2.1 Orthogonal projection

Definition 16 (perpendicular)**.** Let $(V, \langle\ ,\ \rangle)$ be an inner product space. Then, we say a vector v is perpendicular to w if

$$\langle v, w\rangle = 0.$$

We often write $v \perp w$ to indicate two vectors are perpendicular to each other.

Note that the Pythagorean theorem holds under this definition:

$$\boxed{\text{If } \langle v, w\rangle = 0, \text{ then } \|v + w\|^2 = \|v\|^2 + \|w\|^2.}$$

Now, we can define orthogonal projection of x to y.

Definition 17 (Orthogonal projection)**.** Given two vectors $x, y \in (V, \langle\ ,\ \rangle)$ $(y \neq 0)$. $\mathrm{Proj}_y(x)$ is the vector satisfying the following two conditions:

1. $\mathrm{Proj}_y(x)$ is parallel to y.
2. $x - \mathrm{Proj}_y(x) \perp y$.

From this definition, we can assume that $\mathrm{Proj}_y(x) = \alpha \cdot y$, for some $\alpha \in \mathbb{F}$. Since $x - \mathrm{Proj}_y(x) \perp y$, we have

$$\langle x - \alpha \cdot y, y\rangle = 0 \iff \alpha = \frac{\langle x, y\rangle}{\langle y, y\rangle}.$$

We conclude that

$$\boxed{\mathrm{Proj}_y(x) = \frac{\langle x, y\rangle}{\|y\|^2} \cdot y.}$$

Lemma 1. Let $x, y \in (V, \langle\ ,\ \rangle)$ $(y \neq 0)$. Then,

$$\left\|\mathrm{Proj}_y(x)\right\| \leq \|x\|.$$

Moreover, the equality holds if and only if x is parallel to y.

Proof. It follows from the Pythagorean theorem. $\square$

Corollary. $|\langle x, y\rangle| \leq \|x\| \|y\|$, holds for all $x, y \in V$.

It immediate follows from Lemma 1. This inequality is known as "Cauchy's inequality".

Corollary. $\|x + y\| \leq \|x\| + \|y\|$, holds for all $x, y \in V$.

Proof. It is equivalent to prove $\|x + y\|^2 \leq (\|x\| + \|y\|)^2$.

$$\begin{aligned}
& \|x+y\|^2 \leq (\|x\| + \|y\|)^2 \\
\iff \quad & \langle x+y, x+y\rangle \leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 \\
\iff \quad & \|x\|^2 + \langle x, y\rangle + \langle y, x\rangle + \|y\|^2 \leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 \\
\iff \quad & \Re\langle x, y\rangle \leq \|x\| \cdot \|y\|.
\end{aligned}$$

Note that $\Re\langle x, y\rangle \leq |\langle x, y\rangle| \leq \|x\| \cdot \|y\|$. This proves the corollary. $\square$

In general, if we were given a subspace $W \subset V$, we can discuss about $\mathrm{Proj}_W(x)$, the orthogonal projection of x to W.

Definition 18 (Generalization of orthogonal projection)**.** Let W be a subspace of V and let x be a vector in V. Then, $\mathrm{Proj}_W(x)$ is the vector satisfying the following two conditions:

1. $\mathrm{Proj}_W(x) \in W$.
2. $x - \mathrm{Proj}_W(x) \perp W$. That is, $x - \mathrm{Proj}_W(x)$ is perpendicular to any vectors in W.

The existence of $\mathrm{Proj}_W(x)$ in a finite dimensional vector space V follows from the following theorem.

Theorem 19. *Let V be a finite dimensional inner product space and let W be a subspace of V. Define $W^\perp$ as*

$$W^\perp := \{v \in V : \langle v, w\rangle = 0,\ \textit{for all } w \in W\}.$$

Then, $W^\perp$ is a subspace. Moreover, $V = W \oplus W^\perp$.

Proof. It is easy to see that $W^\perp$ is a subspace of V. Recall Theorem 12, we have an isomorphism:

$$\begin{aligned}
V &\simeq V^\vee \\
v &\mapsto l_v(x) = \langle x, v\rangle.
\end{aligned}$$

Note that the image of $W^\perp$ under this map is

$$\{l \in V^\vee : W \subset \ker l\}.$$

By Theorem 7, we have

$$W^{\perp} \simeq (V/W)^{\vee}.$$

Thus,

$$\begin{aligned}\dim_{\mathbb{F}} V &= \dim_{\mathbb{F}} W + (\dim_{\mathbb{F}} V - \dim_{\mathbb{F}} W) \\ &= \dim_{\mathbb{F}} W + \dim_{\mathbb{F}} V/W \\ &= \dim_{\mathbb{F}} W + \dim_{\mathbb{F}} W^{\perp}.\end{aligned}$$

We claim that $W \cap W^{\perp} = \{0\}$. Suppose $x \in W \cap W^{\perp}$, then $\langle x, x\rangle = 0$. This shows that x must be 0. We conclude that

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

□

If we are given a subspace $W \subset V$ and a vector x, then according to Theorem 19, there exist unique vectors $w_x \in W$, $w'_x \in W^{\perp}$ such that

$$x = w_x + w'_x.$$

We define $\mathrm{Proj}_w(x) := w_x$. We now discuss a new idea of (external) direct sum of vector spaces.

Definition 20 (direct sum)**.** Let V_1, V_2 be two vector spaces. Define

$$V_1 \oplus V_2 := \{(v_1, v_2) \in V_1 \times V_2\}.$$

This space has a natural linear structure:

$$\begin{aligned}(v_1, v_2) + (v'_1, v'_2) &:= (v_1 + v'_1, v_2 + v'_2) \\ c(v_1, v_2) &:= (c \cdot v_1, c \cdot v_2)\end{aligned}$$

We shall say $V_1 \oplus V_2$ is the external direct sum of V_1 and V_2.

We can check that:

> If W_1, W_2 are two subspaces of V, such that $W_1 \cap W_2 = \{0\}$. Then,
>
> $$W_1 \oplus_{\text{in}} W_2 \simeq W_1 \oplus_{\text{out}} W_2,$$
>
> where $\oplus_{\text{in}}$ is the original (internal) direct sum.

2.2 Orthonormal basis and Gram-Schimdt process

Definition 21 (orthonormal basis)**.** A set of vectors $\{v_\alpha : \alpha \in \Lambda\}$ is an orthonormal set if $\langle v_\alpha, v_\beta\rangle = 0$ whenever $\alpha \neq \beta$, and $\|v_\alpha\| = 1$ for all $\alpha \in \Lambda$. An orthonormal basis is an orthonormal set which is a basis.

Lemma 2. If $\{v_1, v_2, \ldots, v_r\}$ is an orthonormal set, then it is linearly independent.

Proof. Suppose there exist $\alpha_i \in \mathbb{F}$ such that

$$\sum_{i=1}^{r} \alpha_i \cdot v_i = 0.$$

Then,

$$0 = \langle 0, v_i \rangle = \left\langle \sum_{i=1}^{r} \alpha_i \cdot v_i, v_i \right\rangle = \alpha_i.$$

This completes the proof. □

Remark.

1. If $\dim_{\mathbb{F}} V < \infty$, then any orthonormal set of cardinality equal to n is an orthonormal basis.

2. Let $\mathcal{A}$ be an orthonormal basis. Then, $\Omega = I_n$, where Ω is the matrix of $\langle\ ,\ \rangle$ associated with $\mathcal{A}$.

The existence of orthonormal bases in a finite dimensional inner product space follows from the next theorem. The technique to find such a basis is known as Gram-Schmidt process.

Theorem 22 (Gram-Schmidt process)**.** *Suppose $\{v_1, v_2, \ldots, v_r\}$ is linearly independent. Then, there exists an orthonormal set $\{w_1, w_2, \ldots, w_r\}$ such that*

$$\operatorname{span}_{\mathbb{F}}\{w_1, w_2, \ldots, w_r\} = \operatorname{span}_{\mathbb{F}}\{v_1, v_2, \ldots, v_r\}.$$

Proof. Define u_i and w_i recursively as:

$$\begin{aligned}
u_1 &= v_1 & w_1 &= \frac{u_1}{\|u_1\|} \\
u_2 &= v_2 - \langle v_2, w_1 \rangle \cdot w_1 & w_2 &= \frac{u_2}{\|u_2\|} \\
u_3 &= v_3 - \langle v_3, w_2 \rangle \cdot w_2 - \langle v_3, w_1 \rangle \cdot w_1 & w_3 &= \frac{u_3}{\|u_3\|} \\
&\vdots & &\vdots \\
u_k &= v_k - \sum_{i=1}^{k} \langle v_k, w_i \rangle \cdot w_i & w_k &= \frac{u_k}{\|u_k\|} \\
&\vdots & &\vdots
\end{aligned}$$

We claim that $\operatorname{span}_{\mathbb{F}}\{v_1, \ldots, v_k\} = \operatorname{span}_{\mathbb{F}}\{w_1, \ldots, w_k\}$ and $\{w_1, \ldots, w_k\}$ is an orthonormal set, for each $1 \leq k \leq r$. It is trivial when $k = 1$. Suppose this assertion is true for some $k = m < r$, then $\langle u_{m+1}, w_i \rangle = \langle v_{m+1}, w_i \rangle - \langle v_{m+1}, w_i \rangle = 0$ for $i \leq m$. Also, $v_{m+1} \notin \operatorname{span}_{\mathbb{F}}\{w_1, \ldots, w_m\} = \operatorname{span}_{\mathbb{F}}\{v_1, \ldots, v_m\}$, since $\{v_1, v_2, \ldots, v_r\}$ is linearly independent. We thus have $u_{k+1} \neq 0$, this completes the proof by mathematical induction on k. □

Corollary.

1. If $(V, \langle\ ,\ \rangle)$ is a finite dimensional inner product space over $\mathbb{F}$, then an orthonormal basis exists.

2. Let Ω be a positive definite matrix. From the remark of Definition 15, Ω defines an inner product on $V = \mathbb{F}^n$. Let P be an invertible matrix such that $Pe_i = w_i$, where $\{e_1, \ldots, e_n\}$ is the standard basis of V and $\{w_1, \ldots, x_n\}$ is one orthonormal basis of V with respect to the inner product defined by Ω. Then, Theorem 13 asserts

$$I_n = P^{\mathrm{t}} \cdot \Omega \cdot \overline{P} \implies \Omega = P^{-1^{\mathrm{t}}} \cdot \overline{P^{-1}}.$$

Let $Q = P^{-1^{\mathrm{t}}}$, then we conclude

$$\Omega = Q \cdot Q^*.$$

For each positive definite matrix $\Omega \in M_n(\mathbb{F})$, there is an invertible matrix $Q \in M_n(\mathbb{F})$ such that $\Omega = Q \cdot Q^*$.

Recall that in Theorem 19 we have shown the existence of $\mathrm{Proj}_W(x)$ when W is a subspace of finite dimensional vector space V. In fact, we can derive the same result but using a weaker condition.

Theorem 23 (orthogonal projection revisited)**.** *Let $(V, \langle\ ,\ \rangle)$ be an inner product space. (It could be infinite dimensional.) Let $W \subset V$ be a subspace with finite dimension. Then,* $\mathrm{Proj}_W(x)$ *exists uniquely. In fact,*

$$\mathrm{Proj}_W(x) = \sum_{i=1}^{n} \langle x, w_i \rangle \cdot w_i,$$

where $\{w_1, w_2, \ldots, w_n\}$ is an orthogonal basis of W.

Proof. We first show that $\langle x - \mathrm{Proj}_W(x), w \rangle = 0$, for all $w \in W$. Note that

$$\langle x - \mathrm{Proj}_W(x), w_i \rangle = \langle x, w_i \rangle - \langle x, w_i \rangle = 0,$$

for all $1 \le i \le n$. It remains to show $\mathrm{Proj}_W(x)$ is unique. Let $y \in W$ such that $x - y \in W^\perp$, then

$$\begin{aligned}
\|\mathrm{Proj}_W(x) - y\|^2 &= \langle \mathrm{Proj}_W(x) - y, \mathrm{Proj}_W(x) - y \rangle \\
&= \langle \mathrm{Proj}_W(x) - x + x - y, \mathrm{Proj}_W(x) - y \rangle \\
&= \langle \mathrm{Proj}_W(x) - x, \mathrm{Proj}_W(x) - y \rangle + \langle x - y, \mathrm{Proj}_W(x) - y \rangle \\
&= 0 + 0 = 0\,.
\end{aligned}$$

□

We now generalize the idea of orthogonal projection to the case when the subspace W is not given.

Definition 24 (projection)**.** Let V be an inner product space over $\mathbb{F}$, and let $T : V \to V$ be a linear transformation.

1. We say T is a projection if $T^2 = T$.

2. We say T is an orthogonal projection if $T^2 = T$ and $(\mathrm{Im}T)^\perp = \ker T$.

Remark. Let $T : V \to V$ be an orthogonal projection defined as above. Then, $T(v) = \mathrm{Proj}_W(v)$, where $W := \mathrm{Im}T$.

2.3 Hilbert space

In the previous text, lots of properties of inner product spaces only hold when the space is finite dimensional. This subsection we shall introduce a kind of inner product space that act like a finite dimensional inner product space.

Definition 25 (Hilbert space)**.** Let $(V, \langle\ ,\ \rangle)$ be an inner product space. The norm $\|\cdot\|$ induces a metric d on V. V is said to be a Hilbert space, if (V, d) is a complete metric space in the sense that every Cauchy sequence converges. A subspace $W \subset V$ is closed if W is a Hilbert subspace.

Remark. In analysis, "closedness" of a subspace W means that every convergent sequence in W converges to a point in W. This definition coincides the above definition.

Theorem 26 (existence of orthogonal projection)**.** *Let $(V, \langle\ ,\ \rangle)$ be a Hilbert space and let $W \subset V$ be a closed subset. Then,* $\mathrm{Proj}_W(x)$ *exists uniquely.*

Proof. Let $d := \inf_{w \in W} \|w - x\|$. We claim that there exist a vector $y_0 \in W$ such that $\|y_0 - x\| = d$. By the definition of infimum, there exist y_n such that

$$d \leq \|y_n - x\| < d + \frac{1}{n}.$$

We first show that (y_n) is a Cauchy sequence. Given $\epsilon > 0$. Let $N \in \mathbb{N}$ large enough so that

$$\frac{8d}{N} + \frac{4}{N^2} < \epsilon.$$

By the parallelogram law, we have

$$\begin{aligned}
\|y_n - y_m\|^2 &= 2(\|y_n - x\|^2 + \|y_m - x\|^2) - \|y_n + y_m - 2x\|^2 \\
&< 2\left(\left(d + \frac{1}{n}\right)^2 + \left(d + \frac{1}{m}\right)\right) - 4\left\|\frac{y_n + y_m}{2} - x\right\|^2 \\
&< 4\left(d + \frac{1}{N}\right)^2 - 4d^2 = \frac{8d}{N} + \frac{4}{N^2} < \epsilon,
\end{aligned}$$

where $n, m \geq N$. Hence, (y_n) is a Cauchy sequence. Suppose $y_n \to y_0$, then $\|y_0 - x\| = d$. We now show that $p = x - y_0 \in W^\perp$. Let us introduce two parameters $t \in \mathbb{F}$ and $w \in W$, then we have

$$\begin{aligned}
& \|p - t \cdot w\|^2 = \|x - y_0 - t \cdot w\|^2 \geq d^2 \\
\implies \quad & \|p\|^2 + t^2 \cdot \|w\|^2 - 2\Re\big(\bar{t} \cdot \langle p, w\rangle\big) \geq d^2 \qquad (3) \\
\implies \quad & t^2 \cdot \|w\|^2 - 2\Re\big(\bar{t} \cdot \langle p, w\rangle\big) \geq 0.
\end{aligned}$$

If $\langle p, w\rangle \neq 0$, then $\langle p, w\rangle = r \cdot \exp(i\theta)$ for some $r > 0$. We plug in $t = \epsilon \cdot \exp(i\theta)$ to (3), for small enough $\epsilon > 0$. Then,

$$\epsilon^2 \|w\|^2 \geq 2 \cdot \Re(\epsilon r),$$

which fail to be true when ϵ is small enough. Therefore, $y_0 = \lim y_n = \mathrm{Proj}_W(x)$. □

Next, we introduce the concept of bounded linear functional.

Definition 27 (bounded linear functional)**.** Let $(V, \langle\ ,\ \rangle)$ be a Hilbert space over $\mathbb{F}$. A linear functional $l : V \to \mathbb{F}$ is said to be bounded if there exists $M > 0$ such that

$$|l(v)| \leq M \cdot \|v\|,$$

for all $v \in V$. The set of all bounded linear functional on V is denoted by $V^{\vee}_{\text{bdd}}$. In fact, we can similarly define the concept of bounded linear transformation.

Remark.

1. Any bounded linear functional is a continuous function, with respect to the norm of V and metric on $\mathbb{F}$.

2. Any finite dimensional inner product space V is a Hilbert space, moreover, $V^{\vee}_{\text{bdd}} = V^{\vee}$.

Theorem 28 (Riesz representation theorem)**.** *Let $(V, \langle\ ,\ \rangle)$ be a Hilbert space, and let $l \in V^{\vee}_{\text{bdd}}$ be a bounded linear functional, then there exist $y \in V$, such that*

$$l(x) = \langle x, y\rangle,$$

for all $x \in V$.

Proof. Let l be a bounded linear functional. Then, $N = \ker l$ is a closed subspace of V. (Recall that the preimage under a continuous function of a closed set is closed.) If N is V, then $l = 0$, and we can take $y = 0$. Now, we assume that $N \subsetneq V$, it follows from Theorem 26 that there exists $v \in N^{\perp}$. (Hence $l(v) \neq 0$.) Consider a function $\alpha(x) = l(x)/l(v)$, for all $x \in V$. Then,

$$\begin{aligned} & l(x) = \alpha(x) \cdot l(v) \\ \Longrightarrow\ & l(x - \alpha \cdot v) = 0 \\ \Longrightarrow\ & x - \alpha \cdot v \in N \\ \Longrightarrow\ & \langle x - \alpha \cdot v, v\rangle = 0 \\ \Longrightarrow\ & \langle x, v\rangle = \alpha \cdot \langle v, v\rangle \\ \Longrightarrow\ & l(x) = \langle x, y\rangle, \text{ where } y = \frac{\overline{l(v)}}{\|v\|^2} \cdot v. \end{aligned}$$

□

2.4 Adjoint linear transformation

Definition 29 (adjoint linear transformation)**.** Let $(V, \langle\ ,\ \rangle)$ and $(W, \langle\ ,\ \rangle)$ be two inner product spaces over $\mathbb{F}$ and let $T : V \to W$ be a linear transformation. We define the adjoint of T is the transformation $T^* : W \to V$ such that:

$$\langle T^*(w), v\rangle = \langle v, T(w)\rangle,$$

for all $v \in V$ and $w \in W$.

We now show that T^* exists uniquely if both V and W are finite dimensional.

Theorem 30. *Let V and W be two finite dimensional inner product spaces and let $T : V \to W$ be a linear transformation. Then, T^* exists uniquely.*

Proof. By Theorem 22, there exist orthonormal bases of V and W, say $\mathcal{A} = \{v_1, \ldots, v_n\}$ and $\mathcal{B} = \{w_1, \ldots, w_m\}$, respectively. Let $[T]_{\mathcal{A},\mathcal{B}} = A = (a_{ij})_{m\times n}$. We now assume T^* exists, and let $[T^*]_{\mathcal{B},\mathcal{A}} = (b_{ij})_{n\times m}$. Then,

$$\begin{aligned} & \langle T^*(w_i), v_j\rangle = \langle w_i, T(v_j)\rangle \\ \Longrightarrow\quad & \left\langle \sum_{k=1}^{n} b_{ki}\cdot v_k, v_j\right\rangle = \left\langle w_i, \sum_{l=1}^{m} a_{lj}\cdot w_l\right\rangle \\ \Longrightarrow\quad & b_{ji} = \overline{a_{ij}}. \end{aligned}$$

This shows the uniqueness of T^*. In fact, this also shows the existence of T^*, since we can define:

$$\begin{aligned} T^* : W &\to V \\ [w]_{\mathcal{B}} &\mapsto A^*\cdot [w]_{\mathcal{B}}, \end{aligned}$$

where $[w]_{\mathcal{B}}$ denote the coordinate vector of w with respect to the basis $\mathcal{B}$. The calculations above implies T^* meets the condition of adjoint linear transformation. $\square$

Theorem 31. *Let V, W be inner product spaces over $\mathbb{F}$, and let T_1, T_2 and T be linear transformations from V to W. Suppose T_1^*, T_2^* and $T6*$ exist. Then, the following properties hold:*

1. $(T_1 + T_2)^* = T_1^* + T_2^*$.
2. $(\alpha\cdot T)^* = \overline{\alpha}\cdot T^*$, *for* $\alpha\in\mathbb{F}$.
3. *Let U be an inner product space and let $S : W \to U$ be a linear transformation with the adjoint exists. Then,* $(S\circ T)^* = T^*\circ S^*$.
4. $T^{**} = T$.

The proof is very straightforward, so we omit it.

Theorem 32. *Let $T : V \to W$ be a linear transformation between two "finite dimensional" inner product spaces. Then,*

1. $(\mathrm{Im}T)^\perp = \ker(T^*)$.
2. $(\ker T)^\perp = \mathrm{Im}(T^*)$.

Proof. To show the first assertion, suppose $w \in (\mathrm{Im}T)^\perp$, namely,

$$\begin{aligned} & \langle w, T(v)\rangle = 0, \text{ for all } v\in V. \\ \Longleftrightarrow\quad & \langle T^*(w), v\rangle = 0, \text{ for all } v\in V. \\ \Longleftrightarrow\quad & T^*(w) = 0. \\ \Longleftrightarrow\quad & w\in\ker(T^*). \end{aligned}$$

Similarly, for the second assertion. we assume that $v \in \mathrm{Im}(T^*)$, then $v = T^*(w)$ for some $w \in W$. Note that

$$\langle v, x\rangle = \langle T^*(w), x\rangle = \langle w, T(x)\rangle = 0, \text{ for all } x\in\ker T.$$

Thus, we conclude that $\mathrm{Im}(T^*) \subset (\ker T)^\perp$. By the dimensional formulas, we get $\mathrm{Im}(T^*) = (\ker T)^\perp$. $\square$

Definition 33 (unitary linear transformation (operator))**.** Let $T : V \to W$ be a linear transformation between two inner product spaces (probably infinite dimensional). T is called unitary if

$$\langle T(v_1), T(v_2) \rangle = \langle v_1, v_2 \rangle ,$$

for all $v_1, v_2 \in V$.

The next theorem gives a characterization of unitary operators.

Theorem 34. *Given a linear transformation $T : V \to W$ between two finite dimensional inner product spaces. Then the following statements are equivalent:*

1. *T is unitary.*
2. *$\|T(v)\| = \|v\|$, for all $v \in V$.*
3. *$T^* \circ T = \mathrm{Id}_V$.*
4. *T sends the orthonormal basis to an orthonormal set.*

Proof.
(1) $\Longrightarrow$ (2): Obvious.
(2) $\Longrightarrow$ (1): Consider $\|T(x+y)\|^2 = \|x+y\|^2$.

$$\begin{aligned} &\langle T(x), T(y)\rangle + \langle T(y), T(x)\rangle = \langle x, y\rangle + \langle y, x\rangle \\ \Longrightarrow\ &\Re(\langle T(x), T(y)\rangle) = \Re(\langle x, y\rangle). \end{aligned} \tag{4}$$

If $\mathbb{F} = \mathbb{R}$, then (4) shows that $\langle T(x), T(y)\rangle = \langle x, y\rangle$. If $\mathbb{F} = \mathbb{C}$, then plugging in $y \mapsto i \cdot y$ to equation (4) gives

$$\Re((-i)\cdot \langle T(x), T(y)\rangle) = \Re((-i)\cdot \langle x, y\rangle).$$

Together with equation 4 indicate that T is unitary.
(3) $\Longleftrightarrow$ (1): T is unitary if and only if

$$\begin{aligned} & \langle T(x), T(y)\rangle = \langle x, y\rangle , \text{ for all } x, y \in V \\ \Longleftrightarrow\quad & \langle T^*T(x), y\rangle = \langle x, y\rangle , \text{ for all } x, y \in V \\ \Longleftrightarrow\quad & \langle (T^*T - \mathrm{Id}_V)(x), y\rangle = 0, \text{ for all } x, y \in V \\ \Longleftrightarrow\quad & (T^*T - \mathrm{Id}_V) \equiv 0. \end{aligned}$$

(1) $\Longleftrightarrow$ (4): Let $\mathcal{A} = \{v_1, v_2, \ldots, v_n\}$ be an orthonormal basis of V. Then,

$$\langle T(v_i), T(v_j)\rangle = \langle v_i, v_j\rangle = \begin{cases} 1, & \text{if } i \neq j \\ 0, & \text{if } i = j \end{cases}.$$

Thus, $T(\mathcal{A}) = \{T(v_1), T(v_2), \ldots, T(v_n)\}$ is an orthonormal set.

(4) $\Longleftrightarrow$ (1): Let $x, y \in V$ be two arbitrary vector in V. Let $\mathcal{A} = \{v_1, v_2, \ldots, v_n\}$ be an orthonormal basis of V. Assume

$$x = \sum_{i=1}^{n} \alpha_i \cdot v_i, \quad y = \sum_{i=1}^{n} \beta_i \cdot v_i.$$

Then,

$$\langle T(x), T(y)\rangle = \left\langle T(\sum_{i=1}^{n} \alpha_i \cdot v_i), T(\sum_{i=1}^{n} \beta_i \cdot v_i) \right\rangle = \sum_{i=1}^{n} \alpha_i \cdot \overline{\beta_i} = \langle x, y\rangle .$$

□

2.5 spectral theory of normal operators

Definition 35 (self-adjoint and normal operator)**.** Let $T : V \to V$ be a linear operator on an inner product space V.

1. We say T is self-adjoint, if $T = T*$.

2. We say T is normal, if $T \circ T^* = T^* \circ T$.

Remark. A linear operator $T : V \to V$ is unitary if and only if $T^* = T^{-1}$. (Assume that V is finite dimensional.) Thus, unitary operators and self-adjoint operators are normal.

Theorem 36. *Given $T : V \to V$, a linear operator on finite dimensional space V. The the following statements are equivalent.*

1. *T is normal.*

2. *$\|T(v)\| = \|T^*(v)\|$, for all $v \in V$.*

Proof.
(1) $\Longrightarrow$ (2): Note that

$$\langle T(v), T(v)\rangle = \langle T^*T(v), v\rangle = \langle TT^*(v), v\rangle = \langle T^*(v), T^*(v)\rangle .$$

(2) $\Longrightarrow$ (1): Consider $\|T(x+y)\|^2 = \|T^*(x+y)\|^2$ (and $\|T(x+i\cdot y)\|^2 = \|T^*(x+i\cdot y)\|^2$ if $\mathbb{F} = \mathbb{C}$.) Expanding both equations gives

$$\langle T^*T(x), y\rangle = \langle TT^*(x), y\rangle , \text{ for all } x, y \in V.$$

Thus, $T \circ T^* \equiv T^* \circ T$. □

Corollary. Let $T : V \to V$ be a linear operator on a finite dimensional vector space V. Suppose T is normal, and v is an eigenvector of T with eigenvalue λ. Then, v is an eigenvector of T^* with eigenvalue $\overline{\lambda}$.

Proof. Since T is normal, $S = T - \lambda \cdot \mathrm{Id}_V$ is normal. (In fact, $p(T)$ is normal, for all $p(x) \in \mathbb{F}[x]$.) We have $Sv = 0$. From Theorem 36, we have $\|S^*v\| = \|Sv\| = 0$. Hence, v is in the kernel of $S^* = T^* - \overline{\lambda} \cdot \mathrm{Id}_V$. This completes the proof. □

We now prove an useful lemma.

Lemma 3. Let T be a linear operator on V, such that T^* exists. (We have assumed nothing about whether it is normal.) Then,

$$\ker T^*T = \ker T.$$

Proof. Obviously, $\ker T \subset \ker T^*T$. It suffices to show that $\ker T^*T \subset \ker T$. Let $v \in \ker T^*T$, then,

$$\begin{aligned} T^*T(v) = 0 &\implies \langle T^*T(v), v\rangle = 0 \\ &\implies \langle T(v), T(v)\rangle = 0 \\ &\implies \|T(v)\| = 0 \\ &\implies T(v) = 0. \end{aligned}$$

□

Theorem 37 (Semi-simplicity of normal operators)**.** *Suppose T is a normal operator on V. If $T^n \equiv 0$, for some $n \geq 1$. Then $T \equiv 0$.*

Proof. Let $S = T^*T$. By Lemma 3, it suffices to show $\ker S = V$. Since $T^n = 0$, we have $S^n = 0$. (T^* and T commute.) We may enlarge n so that $n = 2^k$ for some $k \in \mathbb{N}$. Note that

$$\left\| S^{2^{k-1}} v \right\|^2 = \left\langle S^{2^{k-1}} v, S^{2^{k-1}} v \right\rangle = \left\langle \left(S^{2^{k-1}} \right)^* S^{2^{k-1}} v, v \right\rangle = \left\langle S^{2^k} v, v \right\rangle = 0.$$

Repeating this process gives us $S = 0$. $\square$

Before we introduce the next theorem (Theorem 38), we shall first prove another useful result.

Lemma 4. Let V be an inner product space over $\mathbb{F}$, and let $T : V \to V$ be a normal operator on V. Suppose $p(x)$ and $q(x)$ are polynomials in $\mathbb{F}$ with no common roots. Then,

$$\ker(p(T)) \perp \ker(q(T)),$$

that is, $\langle v, w \rangle = 0$, for all $v \in \ker(p(T))$ and $w \in \ker(q(T))$.

Proof. Since p, q have no common roots, there exist $A, B \in \mathbb{F}[x]$, such that

$$A(x)p(x) + B(x)q(x) = 1.$$

Let $v \in \ker(p(T))$ and $w \in \ker(q(T))$. We have $B(T)q(T)(v) = v$. Thus,

$$\langle v, w \rangle = \langle B(T)q(T)(v), w \rangle = \langle q(T)B(T)v, w \rangle = \langle B(T)v, q(T)^*(w) \rangle \overset{(\spadesuit)}{=} \langle B(T)v, 0 \rangle = 0.$$

($\spadesuit$) is true since:

$$\begin{aligned} w \in \ker(q(T)) &\implies \|q(T)(w)\| = 0 \\ &\implies \|q(T)^*(w)\| = 0 \\ &\implies q(T)^*(w) = 0. \end{aligned}$$

$\square$

Theorem 38. *Let $(V, \langle\ ,\ \rangle)$ be an finite dimensional inner product space over $\mathbb{C}$. Let $T : V \to V$ be a normal operator on V. Then, T is diagonalizable. Moreover,*

$$V = \bigoplus_{i=1}^{s} E_{\lambda_i} = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_s}$$

is the orthogonal decomposition of eigenspaces of V. Recall that E_{λ_i} is the eigenspace that which has eigenvalue λ.

Here we give two proofs.

Proof. Let $\mathrm{ch}_T(x)$ be the characteristic polynomial of T. The fundamental theorem of algebra asserts that $\mathrm{ch}_T(x)$ splits completely, that is,

$$\mathrm{ch}_T(x) = \prod_{i=1}^{s} (x - \lambda_i)^{n_i}.$$

Then, we have learnt that $V = \bigoplus_{i=1}^{s} W_i$ in the theory of Jordan forms, where

$$W_i = \ker(T - \lambda_i \cdot \mathrm{Id}_V)^{n_i}.$$

Consider $T|_{W_i}$ on $(W_i, \langle\ ,\ \rangle|_{W_i \times W_i})$. Note that $T|_{W_i}$ is normal and that $(T|_{W_i} - \lambda_i \cdot \mathrm{Id}_{W_i})^{n_i} = 0$. By Theorem 37, we conclude $T|_{W_i} - \lambda_i \cdot \mathrm{Id}_{W_i} = 0$. This implies

$$W_i = \ker(T - \lambda_i \cdot \mathrm{Id}_V)^{n_i} = \ker(T - \lambda_i \cdot \mathrm{Id}_V) = E_{\lambda_i}.$$

It remains to show that each E_{λ_i} is orthogonal to each other. It follows by Lemma 4. $\square$

Here is an alternative proof using mathematical induction.

Proof. Let $\lambda \in \mathbb{C}$ be an eigenvalue of T. Then,

$$E_\lambda = \{v \in V : T(v) = \lambda \cdot v\} \neq \{0\}.$$

Decompose V into $E_\lambda \oplus E_\lambda^\perp$. ($V$ is finite dimensional.) We claim that $E_\lambda^\perp$ is a T-invariant subspace. Let $x \in E_\lambda^\perp$ and $v \in E_\lambda$. Then,

$$\langle T(x), v\rangle = \langle x, T^*(v)\rangle \overset{(\spadesuit)}{=} \langle x, \overline{\lambda} v\rangle = \lambda \langle x, v\rangle = 0.$$

The equality ($\spadesuit$) holds because of Corollary 2.5. On the other hand,

$$\dim E_\lambda^\perp < \dim V.$$

By induction, $T|_{E_\lambda^\perp}$ is diagonalizable and

$$E_\lambda^\perp = \bigoplus_i E_{\lambda_i}.$$

This completes the proof. $\square$

However, Theorem 38 is not true for inner product space over $\mathbb{R}$. But we have the following theorem.

Theorem 39. *Let V be a finite dimensional inner product space over $\mathbb{R}$, and let $T : V \to V$ be a self-adjoint operator on V. Then, T is diagonalizable. Moreover,*

$$V = \bigoplus_{i=1}^{s} E_{\lambda_i},$$

and $E_{\lambda_i} \perp E_{\lambda_j}$ if $i \neq j$.

Proof. In view of the proofs of Theorem 38, it suffices to show that $\mathrm{ch}_T(x)$ splits completely in $\mathbb{R}$. Choose an orthonormal basis $\mathcal{A} = \{v_1, v_2, \ldots, v_n\}$ of V. Define a matrix

$$A := [T]_{\mathcal{A}} = (a_{ij})_{n \times n}.$$

Then, it is well-known that

$$[T^*]_{\mathcal{A}} = A^*.$$

Hence $A^* = A$ since T is self-adjoint. Now, assume $\lambda \in \mathbb{C}$ is an eigenvalue of T. Then, there exists $x \in \mathbb{C}^n \backslash \{0\}$ (column vector) such that

$$Ax = \lambda \cdot x.$$

Consider

$$\overline{\lambda}(x^* \cdot x) = (Ax)^* \cdot x = x^* \cdot A^* \cdot x = x^* \cdot A \cdot x = \lambda \cdot (x^* \cdot x).$$

This indicates

$$\lambda \cdot \|x\|^2 = \overline{\lambda} \cdot \|x\|^2 \implies \lambda \in \mathbb{R}.$$

□

Corollary. Let $A \in M_n(\mathbb{C})$ be a complex normal matrix, that is,

$$A^* \cdot A = A \cdot A^*.$$

Then, there exists an invertible matrix $P \in M_n(\mathbb{C})$ such that:

1. $P \cdot P^* = I_n$.
2. $P^{-1}AP$ is diagonal.

Proof. Let $V = \mathbb{C}^n$ be an inner product space equipped with the standard inner product structure. Let $T : V \to V$ be the operator defined by

$$v \mapsto A \cdot v.$$

Then, the standard basis is orthonormal and hence $A^* = A$ is equivalent to T is self-adjoint. It follows from Theorem 38 that

$$V = \bigoplus_{i=1}^{s} E_{\lambda_i}$$

is a orthogonal decomposition. For each E_{λ_i}, we choose an orthonormal basis

$$\mathcal{A}_i = \{v_{i1}, \ldots, v_{in_i}\}.$$

Then,

$$\mathcal{A} = \bigsqcup_{i=1}^{s} \mathcal{A}_i = \mathcal{A}_1 \sqcup \mathcal{A}_2 \sqcup \cdots \sqcup \mathcal{A}_s$$

is an orthonormal basis. (Because $E_{\lambda_i} \perp E_{\lambda_j}$.) Let P be the matrix sends the standard basis to $\mathcal{A}$. By Theorem 34, we conclude that $P \cdot P^* = P^* \cdot P = I_n$. Also, it is easy to see

$$P^{-1}AP = \begin{pmatrix} \lambda_1 I_{n_1} & 0 & \cdots & 0 \\ 0 & \lambda_2 I_{n_2} & \cdots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_s I_{n_s} \end{pmatrix}.$$

This completes the proof. □

Similarly, one can prove the following result:

Corollary. Let $A \in M_n(\mathbb{R})$ be a real matrix such that $A^{\mathrm{t}} = A$. Then, there exists an invertible matrix $P \in M_n(\mathbb{R})$ such that:

1. $P^{\mathrm{t}} \cdot P = P \cdot P^{\mathrm{t}} = I_n$.
2. $P^{-1}AP$ is diagonal.

In Theorem 39, we show that every self-adjoint operator on vector space over $\mathbb{R}$ is diagonalizable. However, we do not deal with all normal operators. The next theorem is discussing operators over real inner product space.

Theorem 40. *Let $A \in M_n(\mathbb{R})$ be a real normal matrix, that is,*

$$A^{\mathrm{t}} \cdot A = A \cdot A^{\mathrm{t}}.$$

Then, there exists an invertible matrix $P \in M_n(\mathbb{R})$ such that:

1. $P \cdot P^{\mathrm{t}} = P^{\mathrm{t}} \cdot P = I_n$.
2. $P^{-1}AP = (\bigoplus_{i=1}^{s} \lambda_i I_{n_i}) \oplus (\bigoplus_{j=1}^{r} D_j^{m_j})$, *where all $\lambda_i \in \mathbb{R}$, and all D_j have the form:*

$$\begin{pmatrix} \alpha_j & \beta_j \\ -\beta_j & \alpha_j \end{pmatrix}.$$

Remark. Here, we have a little abuse of notation. We write $A \oplus B$ to mean

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix},$$

if both A and B are square matrices.

References

[1] S.H. Friedberg, A.J. Insel, and L.E. Spence. *Linear Algebra*. Featured Titles for Linear Algebra (Advanced) Series. Pearson Education, 2003. ISBN: 9780130084514.

[2] 謝銘倫. 線性代數二. `https://www.youtube.com/playlist?list=PLVJXJebpO4PgW5WgeU_hIml7mSO8Eoggx`. 2016.