Linear Algebra II

陳信睿

2023 年 2 月

Contents

1 Quotient and dual spaces

1.1 Quotient spaces

Definition 1 (Quotient spaces)**.** 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 $\mathcal{B} = \{w_1, w_2, \dots, w_s\}$ be a basis of W and extend $\mathcal{B}$ to $\mathcal{A} = \{w_1, w_2, \dots, w_r\}$ a basis of V. We claim that $\{[w_{s+1}], \dots, [w_s]\}$ is a basis of V/W. To see this, we shall show that:

1. $\{[w_{s+1}], \dots, [w_r]\}$ 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. $\square$

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 spaces

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 $\mathcal{B} = \{v_1, v_2, \dots, v_n\}$ 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 $\mathcal{B}^\vee = \{v_1^\vee, v_2^\vee, \dots, v_n^\vee\}$ 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

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

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 $V^\vee$. □

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 $\mathcal{A} = \{v_1, v_2, \dots, v_n\}$ and $\mathcal{B} = \{w_1, w_2, \dots, w_m\}$ 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,

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

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

Proof. $\square$

2 Inner product spaces

Definition 8 (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 $\langle \cdot, \cdot \rangle$.

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

Definition 9 (norm)**.** For each $v \in V$, define the norm of v as $\|v\| = \langle v, v \rangle^{1/2}$.

Theorem 10 (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

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

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

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

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. □

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 12. *If $\Omega = B \cdot B^*$ for some $B \in M_n(F)$ 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 8. 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

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

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 $\langle x, x\rangle = 0$. □

2.1 Orthogonal projection

Definition 13 (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 14 (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. □

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 $\|x+y\|^2 \le (\|x\|+\|y\|)^2$.

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

Note that $\Re\langle x,y\rangle \le |\langle x,y\rangle| \le \|x\|\cdot\|y\|$. 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 15 (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 $\mathrm{Proj}_W(x)$ follows from the following theorem.

Theorem 16. *Let V be a finite dimensional inner product space and let W be a subspace of V. Define $W^\perp$ as*

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

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 10, 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

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

$\square$