8 General Linear Transformations

8.1 Basic Properties

Definition 8.1 If $T : V \to W$ is a function from a vector space V into a vector space W, then T is called a *linear transformation* from V to W if , for all vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in V and all scalars c,

$$\text{(a) } T(\boldsymbol{u}+\boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v}) \quad \text{(b) } T(c\boldsymbol{u}) = cT(\boldsymbol{u}).$$

In the special case where $V = W$, the linear transformation $T : V \to V$ is called a *linear operator* of V.

Example 8.1 A linear transformation from $\boldsymbol{R}^n$ to $\boldsymbol{R}^m$ is first defined in Definition 2.2 as a function

$$T(x_1, x_2, \ldots, x_n) = (y_1, y_2, \ldots, y_m)$$

for which the equations relating $y_1, y_2, \ldots, y_m$ with $x_1, x_2, \ldots, x_n$ are linear, and it was expressed by a matrix multiplication:

$$T(\boldsymbol{x}) = A\boldsymbol{x}, \text{ where } A = [T(\boldsymbol{e}_1), T(\boldsymbol{e}_2), \ldots, T(\boldsymbol{e}_n)].$$

The matrix A was called the standard matrix and denoted by $A = [T]$ and $T = T_A$. Moreover, linear transformations were characterized by the two properties in Definition 8.1. See Theorem 2.2.

Example 8.2 Let V be an inner product space and W a subspace of V. Then the orthogonal projection $\text{proj}_W : V \to V$ is a linear transformation (or linear operator), and that $\text{proj}_W(V) = W$.

Example 8.3 [Examples 11, 12] Let $C^\infty(a, b)$ be the set of functions that are differentiable for all degrees of differentiation

1. $D : C^\infty(a, b) \to C^\infty(a, b)$ $(f(x) \mapsto f'(x))$ is a linear operator.
2. $I : C^\infty(a, b) \to C^\infty(a, b)$ $(f(x) \mapsto \int_a^x f(t)dt)$ is a linear operator.

Lemma 8.1 (8.1.1) *If $T : V \to W$ is a linear transformation, then*

(a) $T(\boldsymbol{0}) = \boldsymbol{0}$.

(b) $T(-\boldsymbol{v}) = -T(\boldsymbol{v})$ *for all* $\boldsymbol{v} \in V$.

(c) $T\Big(\sum_{i=1}^m k_i\boldsymbol{v}_i\Big) = \sum_{i=1}^m k_iT(\boldsymbol{v}_i)$ *for all* $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m \in V$ *and all scalars* $k_1, k_2, \ldots, k_m$.

Proof. (a): $T(\mathbf{0}) = T(\mathbf{0}+\mathbf{0}) = T(\mathbf{0})+T(\mathbf{0})$. By adding $-T(\mathbf{0})$ on both hand sides, we have $\mathbf{0} = T(\mathbf{0})$.

(b): $T(\boldsymbol{v}) + T(-\boldsymbol{v}) = T(\boldsymbol{v} + (-\boldsymbol{v})) = T(\mathbf{0}) = \mathbf{0}$. Hence by adding $-T(\boldsymbol{v})$ on both hand sides, we have $T(-\boldsymbol{v}) = -T(\boldsymbol{v})$.

(c): This is straightforward as

$$\begin{aligned} T\Big(\sum_{i=1}^{m} k_i\boldsymbol{v}_i\Big) &= T(k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_m\boldsymbol{v}_m) \\ &= T(k_1\boldsymbol{v}_1) + T(k_2\boldsymbol{v}_2) + \cdots + T(k_m\boldsymbol{v}_m) \\ &= k_1T(\boldsymbol{v}_1) + k_2T(\boldsymbol{v}_2) + \cdots + k_mT(\boldsymbol{v}_m) \\ &= \sum_{i=1}^{m} k_iT(\boldsymbol{v}_i). \end{aligned}$$

■

Proposition 8.2 *Let T_1 and T_2 be linear transformations from V to W, and $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ a basis of V[1]. Then the following are equivalent.*

(a) *$T_1 = T_2$, i.e., $T_1(\boldsymbol{v}) = T_2(\boldsymbol{v})$ for all $\boldsymbol{v} \in V$.*

(b) *$T_1(\boldsymbol{v}_i) = T_2(\boldsymbol{v}_i)$ for all $i = 1, 2, \ldots, n$.*

Proof. (a) ⇒ (b) is obvious.

(b) ⇒ (a) follows from Lemma 8.1 (c) as

$$T_1(\boldsymbol{v}) = T_1\Big(\sum_{i=1}^{n} k_i\boldsymbol{v}_i\Big) = \sum_{i=1}^{n} k_iT_1(\boldsymbol{v}_i) = \sum_{i=1}^{n} k_iT_2(\boldsymbol{v}_i) = T_2\Big(\sum_{i=1}^{n} k_i\boldsymbol{v}_i\Big) = T_2(\boldsymbol{v}),$$

when $\boldsymbol{v}$ is expressed as $\boldsymbol{v} = \sum_{i=1}^{n} k_i\boldsymbol{v}_i$. ■

Proposition 8.3 *Let V and W be vector spaces, $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ a basis of V and $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n \in W$. Then there exists a unique linear transformation $T : V \to W$ such that $T(\boldsymbol{v}_i) = \boldsymbol{w}_i$ for all $i = 1, 2, \ldots, n$.*

Proof. The uniqueness follows from Propotion 8.2. Since for every vector $\boldsymbol{v} \in V$ there exist scalars $k_1, k_2, \ldots, k_n$ such that

$$\boldsymbol{v} = k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_n\boldsymbol{v}_n.$$

By Proposition 4.1, $k_1, k_2, \ldots, k_n$ are uniquely determined for each $\boldsymbol{v}$. Let

$$T(\boldsymbol{v}) = k_1\boldsymbol{w}_1 + k_2\boldsymbol{w}_2 + \cdots + k_n\boldsymbol{w}_n.$$

[1] The condition $V = \mathrm{Span}(S)$ is enough.

Then the vector on the right hand side of the equation above is uniquely determined, and this assignment T is a linear transformation satisfying the condition. To see this let $\boldsymbol{u} = \ell_1\boldsymbol{v}_1 + \ell_2\boldsymbol{v}_2 + \cdots + \ell_n\boldsymbol{v}_n \in V$. Then

$$\begin{aligned} T(\boldsymbol{u}+\boldsymbol{v}) &= T(\ell_1\boldsymbol{v}_1 + \ell_2\boldsymbol{v}_2 + \cdots + \ell_n\boldsymbol{v}_n + k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_n\boldsymbol{v}_n) \\ &= T((\ell_1+k_1)\boldsymbol{v}_1 + (\ell_2+k_2)\boldsymbol{v}_2 + \cdots + (\ell_n+k_n)\boldsymbol{v}_n) \\ &= (\ell_1+k_1)\boldsymbol{w}_1 + (\ell_2+k_2)\boldsymbol{w}_2 + \cdots + (\ell_n+k_n)\boldsymbol{w}_n \\ &= \ell_1\boldsymbol{w}_1 + \ell_2\boldsymbol{w}_2 + \cdots + \ell_n\boldsymbol{w}_n + k_1\boldsymbol{w}_1 + k_2\boldsymbol{w}_2 + \cdots + k_n\boldsymbol{w}_n \\ &= T(\ell_1\boldsymbol{v}_1 + \ell_2\boldsymbol{v}_2 + \cdots + \ell_n\boldsymbol{v}_n) + T(k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_2 + \cdots + k_n\boldsymbol{v}_n) \\ &= T(\boldsymbol{u}) + T(\boldsymbol{v}). \end{aligned}$$

We can show $T(k\boldsymbol{v}) = kT(\boldsymbol{v})$ similarly. ■

Proposition 8.4 (8.1.2) *Let $T_1 : U \to V$ and $T_2 : V \to W$ be linear transformations. Then the composition of T_2 with T_1 defined by*

$$T_2 \circ T_1 : U \to W \ (\boldsymbol{x} \mapsto T_2(T_1(\boldsymbol{x}))).$$

is a linear transformation.

Proof. It suffices to prove the conditions (a) and (b) in Definition 8.1.

(a): For $\boldsymbol{u}_1, \boldsymbol{u}_2 \in U$,

$$\begin{aligned} T_2 \circ T_1(\boldsymbol{u}_1 + \boldsymbol{u}_2) &= T_2(T_1(\boldsymbol{u}_1 + \boldsymbol{u}_2)) = T_2(T_1(\boldsymbol{u}_1) + T_1(\boldsymbol{u}_2)) \\ &= T_2(T_1(\boldsymbol{u}_1)) + T_2(T_1(\boldsymbol{u}_2)) = T_2 \circ T_1(\boldsymbol{u}_1) + T_2 \circ T_1(\boldsymbol{u}_2). \end{aligned}$$

(b): For $\boldsymbol{u} \in U$ and a scalary k,

$$T_2 \circ T_1(k\boldsymbol{u}) = T_2(T_1(k\boldsymbol{u})) = T_2(kT_1(\boldsymbol{u})) = kT_2(T_1(\boldsymbol{u})) = k(T_2 \circ T_1(\boldsymbol{u})).$$

Hence $T_2 \circ T_1 : U \to W$ is a linear transformation. ■

8.2 Kernel and Range

Proposition 8.5 (8.2.1) *If $T : V \to W$ is a linear transformation, then*

(a) $\{\boldsymbol{v} \in V \mid T(\boldsymbol{v}) = \boldsymbol{0}\}$ *is a subspace of V.*

(b) $\{T(\boldsymbol{v}) \mid \boldsymbol{v} \in V\}$ *is a subspace of W.*

Proof. We apply Theorem 3.2.

(a): Let $U = \{\boldsymbol{v} \in V \mid T(\boldsymbol{v}) = \boldsymbol{0}\}$. Note that for $\boldsymbol{v} \in V$, $\boldsymbol{v} \in U \Leftrightarrow T(\boldsymbol{v}) = \boldsymbol{0}$. Let $\boldsymbol{v}_1, \boldsymbol{v}_2 \in U$ and k a scalar. Since $T(\boldsymbol{v}_1) = T(\boldsymbol{v}_2) = \boldsymbol{0}$,

$$T(\boldsymbol{v}_1 + \boldsymbol{v}_2) = T(\boldsymbol{v}_1) + T(\boldsymbol{v}_2) = \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}.$$

Hence $\boldsymbol{v}_1 + \boldsymbol{v}_2 \in U$ whenever $\boldsymbol{v}_1, \boldsymbol{v}_2 \in U$. Similarly

$$T(k\boldsymbol{v}_1) = kT(\boldsymbol{v}_1) = k\boldsymbol{0} = \boldsymbol{0}.$$

Hence $k\boldsymbol{v}_1 \in U$ whenever $\boldsymbol{v}_1 \in U$. Therefore U is a subspace of V.

(b): Let $U = \{T(\boldsymbol{v}) \mid \boldsymbol{v} \in V\}$ and $T(\boldsymbol{v}_1), T(\boldsymbol{v}_2) \in U$, where $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$. Then

$$T(\boldsymbol{v}_1) + T(\boldsymbol{v}_2) = T(\boldsymbol{v}_1 + \boldsymbol{v}_2) \in U$$

as $\boldsymbol{v}_1 + \boldsymbol{v}_2 \in V$. Moreover if k is a scalar,

$$kT(\boldsymbol{v}_1) = T(k\boldsymbol{v}_1) \in U$$

and U is a subspace of W. ∎

Definition 8.2 If $T : V \to W$ is a linear transformation, then the set of vectors in V that T maps into $\boldsymbol{0}$ is called the *kernel* of T; it is denoted by $\mathrm{Ker}(T)$. The set of all vectors in W that are images under T of at least one vector in V is called the *range* of T; it is denoted by $\mathrm{Im}(T)$. The dimension of the range of T is called the *rank* of T and is denoted by $\mathrm{rank}(T)$, the dimension of the kernel is called the *nullity* of T and is denoted by $\mathrm{nullity}(T)$.

$\mathrm{Ker}(T) = \{\boldsymbol{v} \in V \mid T(\boldsymbol{v}) = \boldsymbol{0}\} \subset V$, $\mathrm{Im}(T) = \{T(\boldsymbol{v}) \mid \boldsymbol{v} \in V\} \subset W$, and $\mathrm{nullity}(T) = \dim(\mathrm{Ker}(T))$, $\mathrm{rank}(T) = \dim(\mathrm{Im}(T))$.

The following is a generalization of Theorem 5.7. See also Theorem 7.3.

Theorem 8.6 (8.2.3) *If $T : V \to W$ is a linear transformation from an n-dimensional vector space V to a vector space W, then*

$$\mathrm{rank}(T) + \mathrm{nullity}(T) = n.$$

Proof. Let $U = \mathrm{Ker}(T)$ and $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_r\}$ a basis of $\mathrm{Im}(T)$. Hence $r = \dim(\mathrm{Im}(T)) = \mathrm{rank}(T)$. Since $\boldsymbol{w}_i \in \mathrm{Im}(T) = \{T(\boldsymbol{v}) \mid \boldsymbol{v} \in V\}$ for $i = 1, 2, \ldots, r$, there exists $\boldsymbol{v}_i \in V$ such that $f(\boldsymbol{v}_i) = \boldsymbol{w}_i$. Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ be a basis of U. Hence $k = \dim(\mathrm{Ker}(T)) = \mathrm{nullity}(T)$. Our goal is to show $r + k = n = \dim(V)$. It suffices to show that $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_r, \boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is a basis of V.

Linear Independence: Suppose $a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_r\boldsymbol{v}_r + b_1\boldsymbol{u}_1 + b_2\boldsymbol{u}_2 + \cdots + b_k\boldsymbol{u}_k = \boldsymbol{0}$. Since $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k \in \mathrm{Ker}(T)$, by Lemma 8.1

$$\begin{aligned} \boldsymbol{0} &= T(\boldsymbol{0}) = T(a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_r\boldsymbol{v}_r + b_1\boldsymbol{u}_1 + b_2\boldsymbol{u}_2 + \cdots + b_k\boldsymbol{u}_k) \\ &= a_1T(\boldsymbol{v}_1) + a_2T(\boldsymbol{v}_2) + \cdots + a_rT(\boldsymbol{v}_r) \\ &= a_1\boldsymbol{w}_1 + a_2\boldsymbol{w}_2 + \cdots + a_r\boldsymbol{w}_r. \end{aligned}$$

Since $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_r\}$ is a basis of $\mathrm{Im}(T)$, it is linearly independent. Hence $a_1 = a_2 = \cdots = a_r = 0$. Now the first equation yields $b_1\boldsymbol{u}_1 + b_2\boldsymbol{u}_2 + \cdots + b_k\boldsymbol{u}_k = \boldsymbol{0}$ and $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is a basis of U and linearly independent. Hence $b_1 = b_2 = \cdots = b_k = 0$. Therefore, S is linearly independent.

$V = \mathrm{Span}(S)$: Let $\boldsymbol{v} \in V$. Since $T(\boldsymbol{v}) \in \mathrm{Im}(T)$, there exist scalars $a_1, a_2, \ldots, a_r$ such that $T(\boldsymbol{v}) = a_1\boldsymbol{w}_1 + a_2\boldsymbol{w}_2 + \cdots + a_r\boldsymbol{w}_r$. Let $\boldsymbol{w} = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_r\boldsymbol{v}_r$. Then

$$\begin{aligned} T(\boldsymbol{w}) &= T(a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_r\boldsymbol{v}_r) \\ &= a_1T(\boldsymbol{v}_1) + a_2T(\boldsymbol{v}_2) + \cdots + a_rT(\boldsymbol{v}_r) \\ &= a_1\boldsymbol{w}_1 + a_2\boldsymbol{w}_2 + \cdots + a_r\boldsymbol{w}_r \\ &= T(\boldsymbol{v}). \end{aligned}$$

Hence $T(\boldsymbol{v}-\boldsymbol{w}) = T(\boldsymbol{v}) - T(\boldsymbol{w}) = \mathbf{0}$ and $\boldsymbol{v}-\boldsymbol{w} \in \mathrm{Ker}(T)$. Thus there exist scalars $b_1, b_2, \ldots, b_k$ such that $\boldsymbol{v}-\boldsymbol{w} = b_1\boldsymbol{u}_1 + b_2\boldsymbol{u}_2 + \cdots + b_k\boldsymbol{u}_k$ as $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ is a basis of $\mathrm{Ker}(T)$. Therefore

$$\boldsymbol{v} = \boldsymbol{w} + b_1\boldsymbol{u}_1 + b_2\boldsymbol{u}_2 + \cdots + b_k\boldsymbol{u}_k = a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_r\boldsymbol{v}_r + b_1\boldsymbol{u}_1 + b_2\boldsymbol{u}_2 + \cdots + b_k\boldsymbol{u}_k$$

is in $\mathrm{Span}(S)$. This completes the proof. ■

Proposition 8.7 (8.3.1) *If $T : V \to W$ is a linear transformation, then the following are equivalent.*

(a) *T is one-to-one, i.e., injective.*

(b) $\mathrm{Ker}(T) = \{\mathbf{0}\}$.

(c) $\mathrm{nullity}(T) = 0$

Proof. By definition (b) $\Leftrightarrow$ (c). Suppose T is one-to-one. Let $\boldsymbol{v} \in \mathrm{Ker}(T)$. Since $\mathbf{0} \in \mathrm{Ker}(T)$, $T(\mathbf{0}) = \mathbf{0} = T(\boldsymbol{v})$. We have $\boldsymbol{v} = \mathbf{0}$ as T is one-to-one. Hence $\mathrm{Ker}(T) = \{\mathbf{0}\}$. This shows (a) $\Rightarrow$ (b).

Suppose $\mathrm{Ker}(T) = \{\mathbf{0}\}$ and $T(\boldsymbol{v}_1) = T(\boldsymbol{v}_2)$ with $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$. Then $T(\boldsymbol{v}_1 - \boldsymbol{v}_2) = T(\boldsymbol{v}_1) - T(\boldsymbol{v}_2) = \mathbf{0}$ and $\boldsymbol{v}_1 - \boldsymbol{v}_2 \in \mathrm{Ker}(T) = \{\mathbf{0}\}$. Thus $\boldsymbol{v}_1 - \boldsymbol{v}_2 = \mathbf{0}$, or $\boldsymbol{v}_1 = \boldsymbol{v}_2$. Therefore T is one-to-one. ■

Proposition 8.8 (8.3.2) *If V is a finite-dimensional vector space, and $T : V \to V$ is a linear operator, then the following are equivalent.*

(a) *T is one-to-one, i.e., injective.*

(b) $\mathrm{Ker}(T) = \{\mathbf{0}\}$.

(c) $\mathrm{nullity}(T) = 0$.

(d) *The range of T is V, i.e., surjective.*

(e) $\mathrm{rank}(T) = \dim V$.

Proof. (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c) are already shown in Proposition 8.7. Let $n = \dim V$. then by Theorem 8.6, $n = \mathrm{rank}(T) + \mathrm{nullity}(T)$. Hence $\mathrm{nullity}(T) = 0$ if and only if $\mathrm{rank}(T) = n$. Since $\mathrm{rank}(T) = \dim(\mathrm{Im}(T))$ and $\mathrm{Im}(T)$ is a subspace of V by Proposition 8.5, $\mathrm{rank}(T) = n = \dim V$ if and only if $\mathrm{Im}(T) = V$ by Theorem 4.8 (d). This establishes the equivalence. ■

Exercise 8.1 [Quiz 8] Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ and $\boldsymbol{u}$ be vectors in $\boldsymbol{R}^3$ given below.

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -3 \\ -2 \end{bmatrix}, \ \boldsymbol{v}_2 = \begin{bmatrix} -2 \\ 7 \\ 4 \end{bmatrix}, \ \boldsymbol{v}_3 = \begin{bmatrix} 3 \\ -8 \\ -6 \end{bmatrix}, \ \boldsymbol{u} = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}.$$

For $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{R}^3$, let $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^T\boldsymbol{v}$ be the inner product, $U = \mathrm{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$, and $T = \mathrm{proj}_U$. You may quote the facts shown in previous quizzes.

1. Show that $T(\boldsymbol{v}_1) = \boldsymbol{v}_1$, $T(\boldsymbol{v}_2) = \boldsymbol{v}_2$, $T(\boldsymbol{v}_3) = \boldsymbol{v}_3$ and $T(\boldsymbol{u}) = \mathbf{0}$.
2. Show that T is a linear transformation using the definition of linear transformations.
3. Show that $T \circ T = T$.
4. Find $\mathrm{Ker}(T)$, $\mathrm{nullity}(T)$, $\mathrm{Im}(T)$ and $\mathrm{rank}(T)$.
5. Show that there is no linear transformation $T' : U \to U$ such that $T'(\boldsymbol{v}_1) = \boldsymbol{v}_2$, $T(\boldsymbol{v}_2) = \boldsymbol{v}_3$ and $T(\boldsymbol{v}_3) = \boldsymbol{v}_1$.