Chapter 4

Linear Transformations

4.1 Definitions and Examples

4.1.1 Linear Transformation

Definition 4.1.1. *A mapping L from a vector space V into a vector space W (written as $L : V \to W$) is called a* ***linear transformation*** *if for all $\mathbf{v}_1, \mathbf{v}_2 \in V$ and for any scalars α so that*

$$\begin{aligned} &L(\mathbf{v}_1 + \mathbf{v}_2) = L(\mathbf{v}_1) + L(\mathbf{v}_2), \text{ and} \\ &L(\alpha \mathbf{v}) = \alpha L(\mathbf{v}). \end{aligned} \tag{4.1}$$

Remark. The last two conditions can be replaced by one:

$$L(\alpha \mathbf{v}_1 + \beta \mathbf{v}_2) = \alpha L(\mathbf{v}_1) + \beta L(\mathbf{v}_2), \tag{4.2}$$

for all $\mathbf{v}_1, \mathbf{v}_2 \in V$ and for all scalars $\alpha, \beta \in F$, where V, W are vector spaces over the same field F ($F = \mathbb{R}$ or $\mathbb{C}$ in this course).

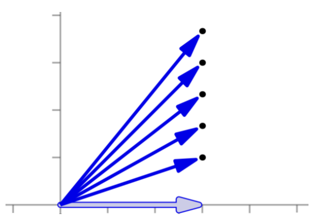
Proof. Prove the equivalence of conditions (4.1) and (4.2). $\square$

Example 4.1.2 (Projection)**.** Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be defined as $L\left[\begin{pmatrix} x \\ y \end{pmatrix}\right] = \begin{pmatrix} x \\ 0 \end{pmatrix}$.

For any $\alpha, \beta \in \mathbb{R}$ and any $\begin{pmatrix} x_1 & y_1 \end{pmatrix}^T, \begin{pmatrix} x_2 & y_2 \end{pmatrix}^T \in \mathbb{R}^2$,

$$
\begin{aligned}
L\left[\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right] &= L\left[\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix}\right] = \begin{pmatrix} x_1 + x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} x_2 \\ 0 \end{pmatrix} \\
&= L\left[\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right] + L\left[\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right] \\
L\left[\alpha \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right] &= L\left[\begin{pmatrix} \alpha x_1 \\ \alpha y_1 \end{pmatrix}\right] = \begin{pmatrix} \alpha x_1 \\ 0 \end{pmatrix} = \alpha \begin{pmatrix} x_1 \\ 0 \end{pmatrix} = \alpha L\left[\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right],
\end{aligned}
$$

so L is a linear transformation. ⚭



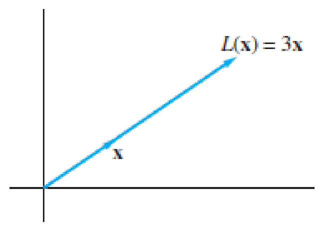


Figure 4.1: Left: Projection Right: Scaling

Example 4.1.3. Let $L : V \to W$ be *zero mapping* defined by $L(\mathbf{v}) = \mathbf{0}_W$ for any $\mathbf{v} \in V$. Then L is a linear transformation. ⚭

Example 4.1.4. Let $L : V \to V$ be *identity mapping* defined by $L(\mathbf{v}) = \mathbf{v}$ for any $\mathbf{v} \in V$. Then L is a linear transformation. ⚭

Definition 4.1.5. *When $W = V$, a linear transformation $L : V \to V$ is called a **linear operator** on V.*

Example 4.1.6 (Scaling)**.** Let L be the mapping defined by $L(\mathbf{x}) = 3\mathbf{x}$ for each

$\mathbf{x} \in \mathbb{R}^2$. Since for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ and $\alpha \in \mathbb{R}$,

$$L(\alpha\mathbf{x}) = 3(\alpha\mathbf{x}) = \alpha(3\mathbf{x}) = \alpha L(\mathbf{x})$$

and

$$L(\mathbf{x}+\mathbf{y}) = 3(\mathbf{x}+\mathbf{y}) = 3\mathbf{x}+3\mathbf{y} = L(\mathbf{x})+L(\mathbf{y})$$

it follows that L is a linear operator. ꝏ

Example 4.1.7. The mapping $L : \mathbb{R}^2 \to \mathbb{R}^3$ defined by $L(\mathbf{x}) = (x_2, x_1, x_1 + x_2)^T$ is linear since $\forall \alpha, \beta \in \mathbb{R}$ and any $\mathbf{x} = (x_1, x_2)^T, \mathbf{y} = (y_1, y_2)^T \in \mathbb{R}^2$,

$$L(\alpha\mathbf{x}) = (\alpha x_2, \alpha x_1, \alpha x_1 + \alpha x_2)^T = \alpha L(\mathbf{x})$$

and

$$L(\mathbf{x}+\mathbf{y}) = L\left[\begin{pmatrix} x_1+y_1 \\ x_2+y_2 \end{pmatrix}\right] = \begin{pmatrix} x_2+y_2 \\ x_1+y_1 \\ x_1+y_1+x_2+y_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 \\ x_1+x_2 \end{pmatrix} + \begin{pmatrix} y_2 \\ y_1 \\ y_1+y_2 \end{pmatrix} = L(\mathbf{x})+L(\mathbf{y}).$$

ꝏ

Example 4.1.8. The mapping $\frac{d}{dx} : P_3 \to P_2$ given by $\frac{d}{dx}(a + bx + cx^2) = b + 2cx$ is linear since $\forall \alpha, \beta \in \mathbb{R}$ and any polynomials $p_1(x), p_2(x) \in P_3$,

$$\begin{aligned}
\tfrac{d}{dx}\left[\, \alpha\, p_1(x) + \beta\, p_2(x) \,\right] &= \frac{d}{dx}\left[(\alpha a_1 + \beta a_2) + (\alpha b_1 + \beta b_2)x + (\alpha c_1 + \beta c_2)x^2 \right] \\
&= (\alpha b_1 + \beta b_2) + 2(\alpha c_1 + \beta c_2)x \\
&= (\alpha b_1 + 2\alpha c_1 x) + (\beta b_2 + 2\beta c_2 x) \\
&= \alpha \cdot \frac{d}{dx}\, p_1(x) + \beta \cdot \frac{d}{dx}\, p_2(x)
\end{aligned}$$

ꝏ

Exercise 4.1.9 (Shear). Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by $L[(x, y)'] = (5x + y, x)'$. Verify that L a linear operator. ✓

Exercise 4.1.10 (Translation). A translation by a vector $\mathbf{a}$ is a transformation of the form $L(\mathbf{x}) = \mathbf{x} + \mathbf{a}$. If $\mathbf{a} \neq \mathbf{0}$, prove that L is not a linear operator on $\mathbb{R}^n$. ✓

Theorem 4.1.11. *For any linear transformation $L : V \to W$, we have $L(\mathbf{0}_V) = \mathbf{0}_W$.*

Exercise 4.1.12. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be mapping defined by $T(x, y) = (x + 1, y + 2)$, for all $(x, y) \in \mathbb{R}^2$. Is T linear? (No, $T(0, 0) = (1, 2) \neq (0, 0)$.) ✓

Theorem 4.1.13. *If $L : V \to W$ is a linear transformation, then*

$$L(\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n) = \alpha_1 L(\mathbf{v}_1) + \alpha_2 L(\mathbf{v}_2) + \cdots + \alpha_n L(\mathbf{v}_n)$$

for any $\alpha_i \in F$ and $\mathbf{v}_i \in V$, $i = 1, \cdots, k$.

Exercise 4.1.14. Let $T : \operatorname{span}\{(-1, 0, 1)', (0, 1, 0)'\} \to \mathbb{R}^2$ be linear mapping defined by $T[(-1, 0, 1)'] = (1, 3)'$ and $T[(0, 1, 0)'] = (2, 4)'$. Find $T[(2, 3, -2)']$. (Ans: $(10, -2)'$.) ✓

Exercise 4.1.15. Let $T : P_3 \to P_4$ be integration mapping defined by $T[p(x)] = \int_1^t p(x)\mathrm{d}x$ for all $p(x) \in P_3$, that is,

$$\begin{aligned} T(a_0 + a_1 x + a_2 x^2) &= \int_1^t a_0 + a_1 x + a_2 x^2 \,\mathrm{d}x = a_0 x + a_1 \tfrac{x^2}{2} + a_2 \tfrac{x^3}{3}\Big|_1^t \\ &= -\left(a_0 + \tfrac{a_1}{2} + \tfrac{a_2}{3}\right) + a_0 t + a_1 \tfrac{t^2}{2} + a_2 \tfrac{t^3}{3}. \end{aligned}$$

Verify that T is linear. ✓

4.1.2 Kernel and Image

Definition 4.1.16 (Kernel). *Let $L : V \to W$ be a linear transformation. Then the **kernel** of L is a subset of V given by*

$$\ker(L) = \{\mathbf{v} \in V \,|\, L(\mathbf{v}) = \mathbf{0}_W\}.$$

Definition 4.1.17 (Image)**.** *The **image/range** of L is defined in the same way as for general functions, that is, a subset of W given by*

$$L(V) = \{\mathbf{w} \in W |\ \mathbf{w} = L(\mathbf{v}) \quad \textit{for } \mathbf{v} \in V\}.$$

Example 4.1.18. Consider the projection $L : \mathbb{R}^2 \to \mathbb{R}^2$ given by $L[(x, y)'] = (x, 0)'$. To find the kernel of L, let $(x, y)' \in \mathbb{R}^2$. Solve $L[(x, y)'] = (x, 0)' = (0, 0)'$, we have $x = 0$. The solutions are $(0, \alpha)^T$ where $\alpha \in \mathbb{R}$. So $\ker(L) = \left\{ \begin{pmatrix} 0 \\ \alpha \end{pmatrix} \in \mathbb{R}^2 \middle| \alpha \in \mathbb{R} \right\}$.

To find the image of L, $L(\mathbb{R}^2) = \{L(\mathbf{v}) \in \mathbb{R}^2 |\ \mathbf{v} \in \mathbb{R}^2\} = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} \in \mathbb{R}^2 \middle| x \in \mathbb{R} \right\}$. ꝏ

Exercise 4.1.19. Let $T : V \to W$ be zero mapping defined by $T(\mathbf{v}) = \mathbf{0}_W$ for all $\mathbf{v} \in V$. Then $\ker(T) = V$ and $T(V) = \{\mathbf{0}_W\}$. ✓

Exercise 4.1.20. Let $T : V \to V$ be identity mapping defined by $T(\mathbf{v}) = \mathbf{v}$ for all $\mathbf{v} \in V$. Then $\ker(T) = \{\mathbf{0}_V\}$ and $T(V) = V$. ✓

Example 4.1.21. Suppose the linear operator $\frac{d}{dx} : P_3 \to P_3$ is given by $\frac{d}{dx}(a + bx + cx^2) = b + 2cx$. The kernel is

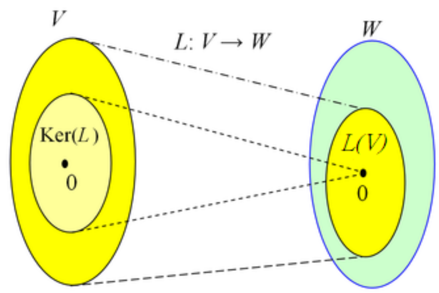
$$\ker(L) = \left\{ a + bx + cx^2 \mid \frac{d}{dx}(a + bx + cx^2) = 0 \right\} = \{a |\ a \in \mathbb{R}\}.$$

The range is

$$\begin{aligned} &\left\{ \frac{d}{dx}(a + bx + cx^2) |\ a, b, c \in \mathbb{R} \right\} \\ =&\{\frac{d}{dx}(a + bx + cx^2) \mid a, b, c \in \mathbb{R}\} \\ =&\{b + 2cx \mid b, c \in \mathbb{R}\}. \end{aligned}$$

ꝏ

Theorem 4.1.22. *Let $L : V \to W$ be a linear transformation. Then the kernel* $\ker(L)$ *is a subspace of V and the image $L(V)$ is a subspace of W.*



Proof. Verify the three defining properties of subspaces. Outline of Proof:

(a)
1. $\ker(L)$ is nonempty, because $\cdots$.
2. $\forall \mathbf{v}_1, \mathbf{v}_2 \in \ker(L)$, then $\cdots$ so $\mathbf{v}_1 + \mathbf{v}_2 \in \ker(L)$.
3. $\forall \mathbf{v}_1 \in \ker(L)$ and $\alpha \in \mathbb{R}$, then $\cdots$ so $\alpha \mathbf{v}_1 \in \ker(L)$.

(b)
1. $L(V)$ is nonempty, because $\cdots$.
2. $\forall \mathbf{v}_1, \mathbf{v}_2 \in L(V)$, then $\cdots$ so $\mathbf{v}_1 + \mathbf{v}_2 \in L(V)$.
3. $\forall \mathbf{v}_1 \in L(V)$ and $\alpha \in \mathbb{R}$, then $\cdots$ so $\alpha \mathbf{v}_1 \in L(V)$.

□

Example 4.1.23. The kernel and range of the linear operator $L(p(x)) = xp'(x)$ on P_3.

For some $p(x) \in P_3$,

$$L(p(x)) = xp'(x) = 0, \quad \to \quad p(x) = \alpha,\ \alpha \in \mathbb{R} \quad \to \quad \ker(L) = P_1.$$

For all $p(x) = a + bx + cx^2 \in P_3$,

$$L(p(x)) = x(b + 2cx) = bx + 2cx^2, \quad \text{where } b, c \in \mathbb{R}. \quad \to \quad L(P_3) = \operatorname{span}\{x, x^2\}.$$

ꝏ

4.2 Matrix Representations of Linear Transformations

4.2.1 from $\mathbb{R}^n$ to $\mathbb{R}^m$

Theorem 4.2.1. *Let A be an $m \times n$ matrix of real numbers. Define $L : \mathbb{R}^n \to \mathbb{R}^m$ by*

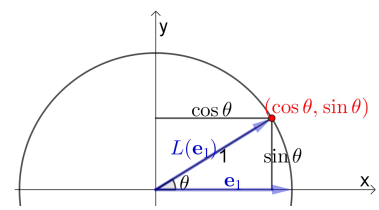
$$L(\mathbf{x}) = A\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n.$$

then L is a linear transformation.

Proof. Verify the definition of linear transformation. □

Example 4.2.2 (Rotation)**.** Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be rotation mapping defined by **counterclockwise** rotating a vector by angle θ, i.e.

$$L\left[\begin{pmatrix} x \\ y \end{pmatrix}\right] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$



L is a linear operator on $\mathbb{R}^2$. ∞

Theorem 4.2.3. *Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then there exists a unique $m \times n$ matrix A such that $L(\mathbf{x}) = A\mathbf{x}$. The matrix given by*

$$A = \begin{bmatrix} L(\mathbf{e}_1) & \cdots & L(\mathbf{e}_n) \end{bmatrix}$$

is called the **matrix representation** *of L and $\{\mathbf{e}_1, \cdots, \mathbf{e}_n\}$ is the standard basis of $\mathbb{R}^n$.*

Proof. Consider $A\mathbf{x}$ as $\sum_i x_i \mathbf{a}_i$, then apply the property of linear transformations. □

Example 4.2.4. Let $L : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation given by $L\left[(x,\ y)^T\right] = (x + y,\ -6x + 3y,\ 2x + 5y)^T$. Since

$$L(\mathbf{e}_1) = [1, -6, 2]^T \quad \text{and} \quad L(\mathbf{e}_2) = [1, 3, 5]^T.$$

So the matrix A representing L is $A = [L(\mathbf{e}_1), L(\mathbf{e}_2)] =$.

∞

Exercise 4.2.5. Let $L : \mathbb{R}^3 \to \mathbb{R}^2$ be a linear mapping given by $L(\mathbf{x}) = \begin{pmatrix} 1x - 3y + 5z \\ 2x - 4y \end{pmatrix}$.
Since $L(\mathbf{e}_1) =$, $L(\mathbf{e}_2) =$, $L(\mathbf{e}_3) =$. Then the matrix representation of L is

$$L(\mathbf{x}) = A\mathbf{x} = [L(\mathbf{e}_1), L(\mathbf{e}_2), L(\mathbf{e}_3)]\mathbf{x} = \begin{pmatrix} 1 & -3 & 5 \\ 2 & -4 & 0 \end{pmatrix} \mathbf{x}.$$

✓

Theorem 4.2.6. *Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and A be its matrix representation. Then*

$$\ker(L) = N(A).$$

Theorem 4.2.7. *Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation and A be its matrix representation. Then*

$$L(\mathbb{R}^n) = Col(A).$$

Example 4.2.8. Let $L : \mathbb{R}^3 \to \mathbb{R}^3$ be given by

$$L(\mathbf{x}) = A\mathbf{x} = \begin{pmatrix} 1 & 2 & 1 \\ -2 & -5 & -4 \\ 3 & 1 & -7 \end{pmatrix} \mathbf{x}.$$

A basis for the kernel of L is $\{(3, -2, 1)^T\}$ since $\ker(L) = \{\alpha(3, -2, 1)^T | \alpha \in \mathbb{R}\}$.
A basis for the range $L(\mathbb{R}^3)$ is $\{(1, -2, 3)^T, (2, -5, 1)^T\}$ since $L(\mathbb{R}^3) = \{\alpha(1, -2, 3)^T + \beta(2, -5, 1)^T : \alpha, \beta \in \mathbb{R}\}$.

∞

Theorem 4.2.9 (Rank-Nullity Theorem)**.** *Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.*

$$\dim(L(\mathbb{R}^n)) + \dim(ker(L)) = \dim(\mathbb{R}^n).$$

Remark. Figure 4.2 shows the geometric effects of some linear mappings on $\mathbb{R}^2$ with matrix representations. The columns of a matrix tell where $L(\mathbf{e}_1)$, $L(\mathbf{e}_2)$ go.

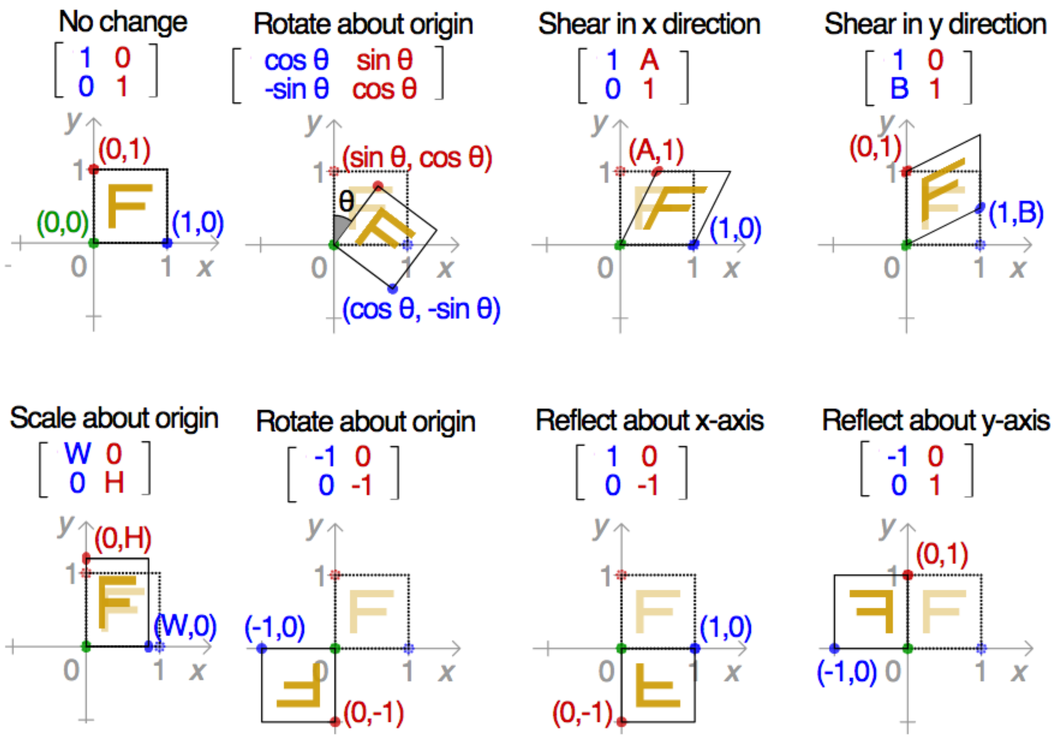


Figure 4.2: Some linear transformations on $\mathbb{R}^2$ with matrix representations

Definition 4.2.10 (Injective). ***A linear transformation* $L : V \longrightarrow W$ *is said to be* **one-to-one** *if* $L(\mathbf{v}_1) = L(\mathbf{v}_2)$ *implies* $\mathbf{v}_1 = \mathbf{v}_2$*, for all* $\mathbf{v}_1, \mathbf{v}_2 \in V$*.*

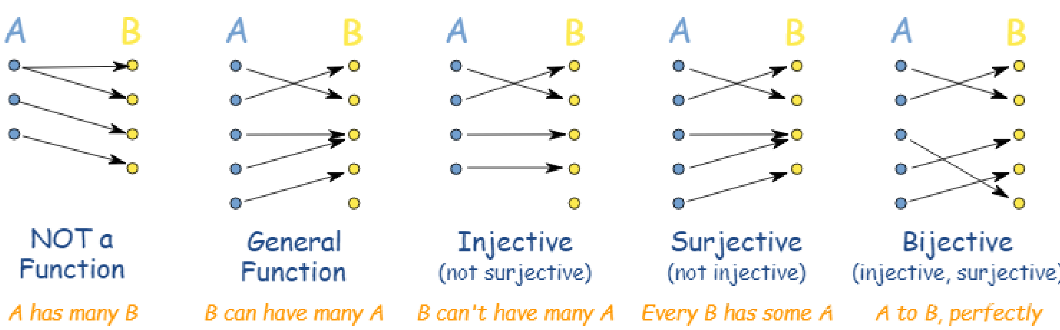


Figure 4.3: Injection, surjection and bijection[1]

[1]Source from the web 'Mathisfun ', https://www.mathsisfun.com/sets/injective-surjective-bijective.html.

Theorem 4.2.11. ** *Let $L : V \to W$ be a linear transformation. Then L is one-to-one (injective) iff $ker(L) = \{\mathbf{0}\}$.*

Proof. Direct proof and proof by contradiction. □

Definition 4.2.12 (Surjective)**.** ***A linear transformation $L : V \longrightarrow W$ is said to be* **onto** *if for any $\mathbf{w} \in W$, there is some $\mathbf{v} \in V$ such that $L(\mathbf{v}) = \mathbf{w}$.*

Corollary 4.2.13. ** *Let $L : V \to W$ be a linear transformation and $\dim V = \dim W$. Then L is injective (one-to-one) iff it is surjective (onto).*

Proof. Apply the previous theorem and the Rank-Nullity theorem. □

Theorem 4.2.14. *Let A be an $m \times n$ matrix with $n \leq m$, and $L : \mathbb{R}^n \to \mathbb{R}^m$ be given by $L(\mathbf{x}) = A\mathbf{x}$. The following are equivalent.*

(a) Columns of A are linearly independent.

(b) The rank of A is n.

(c) The nullity of A is 0.

(d) $A\mathbf{x} = \mathbf{0}$ has only the trivial solution $(0, 0, \cdots, 0)^T$.

(e) The system $A\mathbf{x} = \mathbf{b}$ has at most one solution for every $\mathbf{b} \in \mathbb{R}^m$.

(f) The range of L has dimension n.

(g) $ker(L) = \{\mathbf{0}\}$.

*(h) **L is one-to-one.*

Theorem 4.2.15 ((Equivalent Conditions for Nonsingularity))**.** *Let A be an $n \times n$ matrix. Let $L : \mathbb{R}^n \to \mathbb{R}^n$ be given by $L(\mathbf{x}) = A\mathbf{x}$. The following are equivalent:*

- *A is nonsingular.*
- *$A\mathbf{x} = \mathbf{0}$ has only the trivial solution $\mathbf{0}$.*

- *A is row equivalent to I. (I is the reduced row echelon form of A.)*
- *A can be written as a product of elementary matrices.*
- $\det A \neq 0$.
- *Columns/Rows of A form a basis of $\mathbb{R}^n$.*
- *The rank of A is n.*
- *The nullity of A is 0.*
- *The system $A\mathbf{x} = \mathbf{b}$ has exactly one solution for every $\mathbf{b} \in \mathbb{R}^n$, ($\mathbf{x} = A^{-1}\mathbf{b}$).*
- *The range of L has dimension n.*
- *$ker(L) = \{\mathbf{0}\}$.*
- ***L is one-to-one.*
- *More to go...*

4.2.2 from V to W

We will also see how any linear transformation between finite-dimensional spaces can be represented by a matrix. Here is a series of interesting videos from bilibili.

Theorem 4.2.16 (Matrix Representation Theorem). *Let $L : V \to W$ be a linear transformation with $\dim V = n$ and $\dim W = m$. Let $E = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, $F = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ be ordered bases for V, W, respectively. There exists a unique $m \times n$ matrix A so that*

$$[\mathbf{w}]_F = [L(\mathbf{v})]_F = A[\mathbf{v}]_E, \quad \forall\ \mathbf{v} \in V.$$

In fact,

$$A = \begin{bmatrix} [L(\mathbf{v}_1)]_F & \cdots & [L(\mathbf{v}_n)]_F \end{bmatrix}.$$

We say A is the matrix representation of L with respect to the bases E and F.

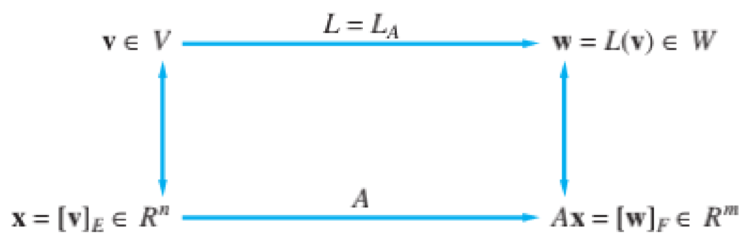


Figure 4.4

Remark. If $V = W$ and $E = F$, then A is called the matrix representing L w.r.t. E.

Proof. Consider $[\mathbf{v}]_E$ and use the fact that the coordinate vector of a linear combination is the linear combination of the coordinate vectors. $\square$

Example 4.2.17. Let $L : \mathbb{R}^2 \to \mathbb{R}^3$ be $L[\mathbf{x}] = (x,\ x+y,\ y)^T$. Find the matrix A representing L with respect to the ordered basis

$$E = \{(2,4)^T, (-1,6)^T\}, \quad \text{and} \quad F = \{(1,1,1)^T, (0,1,0)^T, (0,0,1)^T\}.$$

We have $L(\mathbf{v}_1) = L\left[\begin{pmatrix} 2 \\ 4 \end{pmatrix}\right] = \begin{pmatrix} 2 \\ 6 \\ 4 \end{pmatrix}$ and $L(\mathbf{v}_2) = L\left[\begin{pmatrix} -1 \\ 6 \end{pmatrix}\right] = \begin{pmatrix} -1 \\ 5 \\ 6 \end{pmatrix}$.

<u>Method 1:</u>

$$L(\mathbf{v}_1) = 2\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 4\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 2\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad L(\mathbf{v}_2) = -\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 6\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + 7\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Thus, $A = ([L(\mathbf{v}_1)]_F,\ [L(\mathbf{v}_2)]_F) = \begin{pmatrix} 2 & -1 \\ 4 & 6 \\ 2 & 7 \end{pmatrix}$.

**<u>Method 2:</u> Transform the augmented matrix $(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3 | L(\mathbf{v}_1), L(\mathbf{v}_2))$ to rref:

$$(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3 | L(\mathbf{v}_1), L(\mathbf{v}_2)) = \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 2 & -1 \\ 1 & 1 & 0 & 6 & 5 \\ 1 & 0 & 1 & 4 & 6 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|cc} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 4 & 6 \\ 0 & 0 & 1 & 2 & 7 \end{array}\right) = (I \mid A).$$

∞

Exercise 4.2.18. Let $L : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by $L(\mathbf{x}) = \begin{pmatrix} 1x - 3y + 5z \\ 2x - 4y \end{pmatrix}$. Find the matrix A representing L with respect to the ordered basis $\{(1,0,1)^T, (0,1,1)^T, (0,0,1)^T\}$ and $\{(1,2)^T, (3,5)^T\}$. $A = \left(\begin{smallmatrix} -24 & -22 & -25 \\ 10 & 8 & 10 \end{smallmatrix}\right)$. ✓

Example 4.2.19. Let $L : P_2 \to P_3$ be a linear transformation so that

$$L(1) = 4(1) + 6x + 8x^2, \qquad L(x) = 5(1) + 7x + 9x^2.$$

Denote $E = \{1, x\}$ and $F = \{1, x, x^2\}$, then the matrix A representing L (with respect to the standard bases) is

$$A = \Big[[L(1)]_F, \quad [L(x)]_F\Big] = \begin{pmatrix} 4 & 5 \\ 6 & 7 \\ 8 & 9 \end{pmatrix}.$$

∞

Example 4.2.20. Let $L : P_3 \to P_3$ be a linear operator defined as

$$L(p(t)) = p(1) + p'(1)(t-1).$$

Find the matrix A representing L with respect to the ordered bases $E = \{t^2-1, t+1, 1\}$. Since

$$\begin{aligned} L(t^2-1) &= (1^2-1) + 2(t-1) = 2(t-1) = 2(t+1) - 4 \\ L(t+1) &= (1+1) + 1(t-1) = t+1 \\ L(1) &= 1 + 0 \cdot (t-1) = 1 \end{aligned}$$

Then

$$A = \Big[[L(t^2-1)]_E, \quad [L(t+1)]_E, \quad [L(1)]_E\Big] = \begin{pmatrix} 0 & 0 & 0 \\ 2 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}.$$

∞

4.2.3 ** Composition of Linear Transformations

Theorem 4.2.21 (Composite Linear Transformations)**.** *Let V, W and U be finite dimensional vector spaces with ordered bases α, β and γ respectively. Let $L : V \to W$ and $T : W \to U$ be linear transformations. Then the matrix representation of the composite linear transformation $T \circ L : V \to U$ (or simply TL) is*

$$[T \circ L]_\alpha^\gamma = [T]_\beta^\gamma [L]_\alpha^\beta.$$

Proof. 1. Verify that the composite transformation $T \circ L : V \to U$ is linear by the definition.

2. The matrix representation of the composite linear transformation can be found as a matrix product. □

Example 4.2.22. Let $L : \mathbb{R}^2 \to \mathbb{R}^3$ and $T : \mathbb{R}^3 \to \mathbb{R}^3$ be

$$L\left[\begin{pmatrix} x \\ y \end{pmatrix}\right] = \left[\begin{pmatrix} x \\ x+y \\ y \end{pmatrix}\right], \quad \text{and} \quad T\left[\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right] = \begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix}.$$

To find the matrix representing $T \circ L$ with respect to the ordered basis $E = \{(2,4)^T, (-1,6)^T\}$ and $F = \{(1,1,1)^T, (0,1,0)^T, (0,0,1)^T\}$, we refer to Example 4.2.17 and the matrix representing L w.r.t. the bases E and F is

$$[L]_E^F = \begin{pmatrix} 2 & -1 \\ 4 & 6 \\ 2 & 7 \end{pmatrix}.$$

Further since

$$T((1,1,1)') = (1,1,1)', \quad T((0,1,0)') = (0,1,0)' \quad \text{and} \quad T((0,0,1)') = (1,0,0)',$$

the matrix representation of T w.r.t. the basis F is

$$[T]_F = ([T((1,1,1)')]_F,\ [T((0,1,0)')]_F,\ [T((0,0,1)')]_F) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \end{pmatrix}.$$

Thus, the matrix representing $T \circ L$ w.r.t the ordered basis E and F is

$$[T \circ L]_E^F = [T]_F \, [L]_E^F = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 4 & 6 \\ 2 & 7 \end{pmatrix} = \begin{pmatrix} 4 & 6 \\ 2 & -1 \\ -2 & -7 \end{pmatrix}.$$

ꝏ

Corollary 4.2.23. *Let V be finite dimensional vector space with an ordered basis α. Let $L : V \to V$ and $T : V \to V$ be linear mappings. Then the matrix representation of the composite linear transformation $T \circ L : V \to V$ is*

$$[T \circ L]_\alpha = [T]_\alpha [L]_\alpha,$$

and for all positive integer n, the matrix representation of $\underbrace{L \circ \cdots \circ L}_{n}$ is

$$[L^n]_\alpha = ([L]_\alpha)^n .$$

Example 4.2.24. Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be the reflection about x-axis and $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the reflection about y-axis, i.e.

$$L \left[\begin{pmatrix} x \\ y \end{pmatrix} \right] = \begin{pmatrix} x \\ -y \end{pmatrix}, \quad \text{and} \quad T \left[\begin{pmatrix} x \\ y \end{pmatrix} \right] = \begin{pmatrix} -x \\ y \end{pmatrix}.$$

Then

$$T \circ L \left[\begin{pmatrix} x \\ y \end{pmatrix} \right] = T \begin{pmatrix} x \\ -y \end{pmatrix} = \begin{pmatrix} -x \\ y \end{pmatrix}, \quad \text{and} \quad L \circ T \left[\begin{pmatrix} x \\ y \end{pmatrix} \right] = L \begin{pmatrix} -x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ -y \end{pmatrix}.$$

Under the standard basis $\alpha = \{\mathbf{e}_1, \mathbf{e}_2\}$ of $\mathbb{R}^2$, we have

$$[L]_\alpha = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad [T]_\alpha = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Then

$$[TL]_\alpha = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \text{and} \quad [LT]_\alpha = [L]_\alpha [T]_\alpha = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

It means that the reflection about x-axis then about y-axis is equivalent to the reflection about y-axis first then about x-axis. ꝏ

Remark. The last example is special in the sense that the matrix product $AB = BA$. However, for most cases, $AB \neq BA$.

4.2.4 **Inverse of Linear Transformations

Theorem 4.2.25 (Inverse of Linear Transformations). *A linear transformation $L : U \to V$ is invertible if and only if $\dim U = \dim V$ and any of its matrix representation is nonsingular.*

Corollary 4.2.26. *If a linear mapping $L : U \to V$ is invertible, then the matrix representation of $L^{-1} : V \to U$ equals to the inverse of the matrix representing L, i.e.*

$$[L^{-1}] = [L]^{-1}.$$

Example 4.2.27 (Rotation). Let $L : \mathbb{R}^2 \to \mathbb{R}^2$ be rotation mapping defined by **counterclockwise** rotating a vector by angle θ, i.e.

$$L\left[\begin{pmatrix} x \\ y \end{pmatrix}\right] = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be rotation mapping defined by **counterclockwise** rotating a vector by angle $-\theta$, i.e.

$$T\left[\begin{pmatrix} x \\ y \end{pmatrix}\right] = \begin{pmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Clearly, $T = L^{-1}$. Under the standard basis, their matrices are inverses of each other.

ꝏ

Theorem 4.2.28. *If $L : U \to V$ is invertible if and only if L is bijective.*

Example 4.2.29. Let $L : P_3 \to P_3$ be defined as $L(p(t)) = p'(t)$, for all $p(t) \in P_3$. Under the standard basis $\alpha = \{1, t, t^2\}$ of P_3, we have

$$[L]_\alpha = \Big([L(1)]_\alpha, \quad [L(t)]_\alpha, \quad [L(t^2)]_\alpha\Big) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then

$$[L^3]_\alpha = ([L]_\alpha)^3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

It means that L^3 is a zero mapping which maps all polynomials of degree two to the zero polynomial. Also, by Calculus, when taking derivatives on a degree 2 polynomial for 3 times, the result must be 0, i.e. the derivation cannot be "undone".

$$\frac{\mathrm{d}}{\mathrm{d}t}(a + bt + ct^2) = 0.$$

The differential operator is not invertible, while the matrix $[L]_\alpha$ is also singular. ⚭

4.3 Similarity

Recall in Chapter 3.5, we have used the transition matrix to perform change of bases. Let E, F be two ordered bases for the vector space V of dimension n. Then there exists a unique matrix $[I]_F^E$, namely the **transition matrix** from F to E, such that

$$[\mathbf{x}]_E = [I]_F^E \, [\mathbf{x}]_F, \quad \forall \mathbf{x} \in V.$$

- The transition matrix is nonsingular.
- If $[I]_F^E$ is the transition matrix from F to E, then $([I]_F^E)^{-1}$ is the transition matrix from E to F.

Remark. In $\mathbb{R}^n$, the transition matrix from an ordered basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ to the standard basis is simply

$$S = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_n \end{bmatrix}.$$

Theorem 4.3.1. *Let L be a linear operator on V, and E, F be two ordered bases of V. If A is the matrix representing L with respect to E and B is the matrix representing L with respect to F, then*

$$B = S^{-1}AS,$$

where S is the transition matrix from F to E.

$$\begin{tikzcd}
\text{Basis } E\text{: } V \arrow[r, "L", "A"'] & V \arrow[d, "S^{-1}"'] \\
\text{Basis } F\text{: } V \arrow[u, "S"'] \arrow[r, "B", "L"'] & V
\end{tikzcd}$$

Figure 4.5

Proof. Show $SB = AS$. □

Example 4.3.2. Let $\frac{\mathrm{d}}{\mathrm{d}x}$ be the differential operator on P_3.

(a) Find the matrix A representing $\frac{d}{dx}$ with respect to $E = \{1, x, x^2\}$.

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}x}(1) & = 0(1) + 0x + 0x^2 \\ \frac{\mathrm{d}}{\mathrm{d}x}(x) & = 1(1) + 0x + 0x^2 \\ \frac{\mathrm{d}}{\mathrm{d}x}(x^2) & = 0(1) + 2x + 0x^2 \end{cases} \quad \to \quad A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

(b) Find the matrix B representing $\frac{\mathrm{d}}{\mathrm{d}x}$ with respect to $F = \{1, 2x, 4x^2 - 2\}$.

The transition matrix U from F to E is

$$U = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \qquad \text{then} \qquad B = U^{-1}AU = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix}.$$

Definition 4.3.3. *Two $n \times n$ matrices A and B are said to be* ***similar*** *if there exists a nonsingular matrix S such that $B = S^{-1}AS$.*

From the above theorem: If two matrices A and B are similar, then they represent the same linear transformation (with respect to different bases).

Properties If A and B are similar, then the following statements are true.

1. A^T and B^T are similar.
2. $\det(A) = \det(B)$.
3. A is nonsingular if and only if B is nonsingular. Further A^{-1} and B^{-1} are similar.
4. $\text{rank}(A) = \text{rank}(B)$. Use the fact that $\text{rank}(A) = \text{rank}(PA) = \text{rank}(AQ)$ if P, Q are nonsingular matrices.
5. there exists a nonsingular matrix S such that $B^k = S^{-1}A^kS$.

Proof. 1-2. Exercises.

3. Recall $\det(A) \neq 0$ if and and if A is nonsingular.
4. Recall "If both P and Q are nonsingular, then $\text{rank}(PAQ) = \text{rank}(A)$."
5. Exercise.

$\square$

Exercises from Leon's textbook (9th Ed.):

1. **Sec 4.1 — Q1, Q3-9, Q12-19, Q20**-25****
2. **Sec 4.2 — Q1-8, Q11, Q14, Q16-18, Q20**
3. **Sec 4.3 — Q1-5, Q7-15**