9 Matrices and Linear Transformations

9.1 Matrices and Linear Transformations

Definition 9.1 Suppose that V is an n-dimensional vector space with a basis $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ and W is an m-dimensional vector space with a basis $B' = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$. For $\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n \in V$, the vector $[x_1, x_2, \ldots, x_n]^T \in \boldsymbol{R}^n$ is called the *coordinate vector of* $\boldsymbol{x}$ *with respect to the basis* B and denoted by $[\boldsymbol{x}]_B$. Similarly for $\boldsymbol{y} = y_1\boldsymbol{w}_1 + y_2\boldsymbol{w}_2 + \cdots + y_m\boldsymbol{w}_m$, $[\boldsymbol{y}]_{B'} = [y_1, y_2, \ldots, y_m]^T \in \boldsymbol{R}^m$ is the coordinate vector of $\boldsymbol{y}$ with respect to the basis B'.

Let T be a linear transformation from V to W. Then the $m \times n$ matrix A defined by

$$A = [[T(\boldsymbol{v}_1)]_{B'}, [T(\boldsymbol{v}_2)]_{B'}, \ldots, [T(\boldsymbol{v}_m)]_{B'}]$$

is called the *matrix for* T *with respect to the bases* B *and* B' and denoted by $[T]_{B',B}$.

When $V = W$ and $B = B'$, we write $[T]_B$ for $[T]_{B,B}$ and $[T]_B$ is called the *matrix for* T *with respect to the basis* B.

Proposition 9.1 *Under the notation in Definition 9.1 the following hold.*

(a) $[T]_{B',B}[\boldsymbol{x}]_B = [T(\boldsymbol{x})]_{B'}$.

(b) $[T]_B[\boldsymbol{x}]_B = [T(\boldsymbol{x})]_B$, *with* $V = W$.

Proof. Since (b) is obtained by setting $B' = B$ and $V = W$, it suffices to prove (a).

Since $\boldsymbol{x}' \mapsto [T]_{B',B}(\boldsymbol{x}')$ is a linear mapping from $\boldsymbol{R}^n$ to $\boldsymbol{R}^m$, we compute the both hand sides at basis vectors. Note that $[\boldsymbol{v}_i]_B = \boldsymbol{e}_i$.

$$[T(\boldsymbol{v}_i)]_{B'} = [T]_{B',B}\boldsymbol{e}_i = [T]_{B',B}[\boldsymbol{v}_i]_B.$$

Hence the equality holds for all $\boldsymbol{x}$. ∎

Proposition 9.2 (8.4.2) *Let* $T_1 : U \to V$ *and* $T_2 : V \to W$ *be linear transformations and* B, B' *and* B'' *basis of* U, V, *and* W *respectively. Then*

$$[T_2 \circ T_1]_{B'',B} = [T_2]_{B'',B'}[T_1]_{B',B}.$$

Proof. Let $\boldsymbol{x} \in U$. Then

$$\begin{aligned} [T_2 \circ T_1]_{B'',B}[\boldsymbol{x}]_B &= [(T_2 \circ T_1)(\boldsymbol{x})]_{B''} = [T_2((T_1)(\boldsymbol{x}))]_{B''} = [T_2]_{B'',B'}[T_1(\boldsymbol{x})]_{B'} \\ &= [T_2]_{B'',B'}([T_1]_{B',B}[\boldsymbol{x}]_B) = ([T_2]_{B'',B'}[T_1]_{B',B})[\boldsymbol{x}]_B. \end{aligned}$$

Therefore $[T_2 \circ T_1]_{B'',B} = [T_2]_{B'',B'}[T_1]_{B',B}$. ∎

Proposition 9.3 (8.4.3) *Let* $T : V \to V$ *be a linear transformation. If* B *is a basis of* V, *then the following are equaivalent:*

(a) T *is one-to-one.*

(b) $[T]_B$ *is invertible.*

Morover, when these equivalent conditions hold,

$$[T^{-1}]_B = [T]_B^{-1}.$$

Proof. (a) $\Rightarrow$ (b): Suppose T is one-to-one. Then T is bijective by Proposition 8.8. Hence there is a linear transformation T^{-1} such that $T \circ T^{-1} = T^{-1} \circ T = I$. Then by Proposition 9.2,

$$[T]_B[T^{-1}]_B = [T \circ T^{-1}]_B = [I]_B = [T^{-1} \circ T]_B = [T^{-1}]_B[T]_B.$$

Since $[I]_B = I$, $[T]_B$ is invertible and $[T^{-1}]_B = [T]_B^{-1}$.

(b) $\Rightarrow$ (a): Suppose $[T]_B$ is invertible and $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. Let T' be a linear operator on V defined by $[T'(\boldsymbol{x})]_B = ([T]_B)^{-1}[\boldsymbol{x}]_B$. Then

$$\begin{aligned}
[(T' \circ T)(\boldsymbol{x})]_B &= [T'(T(\boldsymbol{x}))]_B = ([T]_B)^{-1}[T(\boldsymbol{x})]_B \\
&= ([T]_B)^{-1}([T]_B[\boldsymbol{x}]_B) = ([T]_B)^{-1}[T]_B[\boldsymbol{x}]_B = [\boldsymbol{x}]_B, \text{ and} \\
[(T \circ T')(\boldsymbol{x})]_B &= ([T]_B[T']_B)[\boldsymbol{x}]_B = [T]_B[T'(\boldsymbol{x})]_B \\
&= [T]_B([T]_B)^{-1}[\boldsymbol{x}]_B) = [T]_B([T]_B)^{-1}[\boldsymbol{x}]_B = [\boldsymbol{x}]_B
\end{aligned}$$

Therefore $T \circ T' = I = T' \circ T$. ■

9.2 Similarity

Theorem 9.4 (8.5.2) *Let $T : V \to V$ be a linear operator on a finite-dimensional vector space V, and let B and B' be bases for V. Then*

$$[T]_{B'} = P^{-1}[T]_B P$$

where P is the transition matrix from B' to B.

Proof. Let $P = [I]_{B,B'}$. Then $P^{-1} = [I]_{B',B}$. Hence

$$P^{-1}[T]_B P = P^{-1} = [I]_{B',B}[T]_B[I]_{B,B'} = [I \circ T \circ I]_{B',B'} = [T]_{B',B'}.$$

■

Definition 9.2 If A and B are square matrices, we say that B *is similar to* A if there is an invertible matrix P such that $B = P^{-1}AP$.

Example 9.1 Let $V = \boldsymbol{R}^3$. In Exercises we showed that V has three bases.

$$B = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\},\ B' = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{e}_1\},\ \text{and } B'' = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\},$$

where

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -3 \\ -2 \end{bmatrix},\ \boldsymbol{v}_2 = \begin{bmatrix} -2 \\ 7 \\ 4 \end{bmatrix},\ \boldsymbol{v}_3 = \begin{bmatrix} 3 \\ -8 \\ -6 \end{bmatrix}, \boldsymbol{u}_1 = \begin{bmatrix} \frac{1}{\sqrt{14}} \\ \frac{-3}{\sqrt{14}} \\ \frac{-2}{\sqrt{14}} \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} \frac{3}{\sqrt{70}} \\ \frac{5}{\sqrt{70}} \\ \frac{-6}{\sqrt{70}} \end{bmatrix}, \boldsymbol{u}_3 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ 0 \\ \frac{1}{\sqrt{5}} \end{bmatrix}.$$

The first is the standard basis, and the last is an orthonormal basis. Let $T = \mathrm{proj}_U$. We describe $[T]_B$, $[T]_{B'}$, $[T]_{B''}$ and $[I]_{B',B}$, $[I]_{B'',B}$, $I_{B'',B'}$.

$[T]_B$**:** By Quiz 7-5,

$$\boldsymbol{e}_1 = \frac{1}{\sqrt{14}}\boldsymbol{u}_1+\frac{3}{\sqrt{70}}\boldsymbol{u}_2+\frac{2}{\sqrt{5}}\boldsymbol{u}_3, \boldsymbol{e}_2 = \frac{-3}{\sqrt{14}}\boldsymbol{u}_1+\frac{5}{\sqrt{70}}\boldsymbol{u}_2, \boldsymbol{e}_3 = \frac{-2}{\sqrt{14}}\boldsymbol{u}_1-\frac{6}{\sqrt{70}}\boldsymbol{u}_2+\frac{1}{\sqrt{5}}\boldsymbol{u}_3.$$

$$\begin{aligned} T(\boldsymbol{e}_1) &= \frac{1}{\sqrt{14}}\boldsymbol{u}_1 + \frac{3}{\sqrt{70}}\boldsymbol{u}_2 = \boldsymbol{e}_1 - \frac{2}{\sqrt{5}}\boldsymbol{u}_3 = \frac{1}{5}[1,0,-2]^T \\ T(\boldsymbol{e}_2) &= \boldsymbol{e}_2 \\ T(\boldsymbol{e}_3) &= \frac{-2}{\sqrt{14}}\boldsymbol{u}_1 - \frac{6}{\sqrt{70}}\boldsymbol{u}_2 = \boldsymbol{e}_3 - \frac{1}{\sqrt{5}}\boldsymbol{u}_3 = \frac{1}{5}[-2,0,4]^T \end{aligned}$$

Hence

$$[T]_B = \begin{bmatrix} \frac{1}{5} & 0 & -\frac{2}{5} \\ 0 & 1 & 0 \\ -\frac{2}{5} & 0 & \frac{4}{5} \end{bmatrix}$$

$[T]'_B$ Since $T(\boldsymbol{e}_1) = \frac{1}{5}[1,0,-2]^T = \frac{1}{5}(7\boldsymbol{v}_1 + 3\boldsymbol{v}_2)$,

$$[T]_{B'} = \begin{bmatrix} 1 & 0 & \frac{7}{5} \\ 0 & 1 & \frac{3}{5} \\ 0 & 0 & 0 \end{bmatrix}$$

$[T]_{B''}$**:** Since $T(\boldsymbol{u}_1) = \boldsymbol{u}_1$, $T(\boldsymbol{u}_2) = \boldsymbol{u}_2$ and $T(\boldsymbol{u}_3) = \boldsymbol{0}$,

$$[T]_{B''} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$[I]_{B',B}$**,** $[I]_{B'',B}$**,** $[I]_{B'',B'}$**:** We have as follows:

$$\begin{aligned} [I]_{B',B} &= [\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{e}_1] = \begin{bmatrix} 1 & -2 & 1 \\ -3 & 7 & 0 \\ -2 & 4 & 0 \end{bmatrix} \\ [I]_{B'',B} &= [\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3] = \begin{bmatrix} \frac{1}{\sqrt{14}} & \frac{3}{\sqrt{70}} & \frac{2}{\sqrt{5}} \\ \frac{-3}{\sqrt{14}} & \frac{5}{\sqrt{70}} & 0 \\ \frac{-2}{\sqrt{14}} & \frac{-6}{\sqrt{70}} & \frac{1}{\sqrt{5}} \end{bmatrix} \\ [I]_{B'',B'} &= \begin{bmatrix} \frac{\sqrt{14}}{14} & \frac{31\sqrt{70}}{70} & \frac{-7\sqrt{5}}{10} \\ 0 & \frac{\sqrt{70}}{5} & \frac{-3\sqrt{5}}{10} \\ 0 & 0 & \frac{\sqrt{5}}{2} \end{bmatrix} \end{aligned}$$

Recall the situation in Definition 9.1. Suppose

$$T(\boldsymbol{v}_i) = \sum_{j=1}^{m} a_{j,i}\boldsymbol{w}_j = a_{1,i}\boldsymbol{w}_1 + a_{2,i}\boldsymbol{w}_2 + \cdots + a_{m,i}\boldsymbol{w}_m.$$

Then $[T(\boldsymbol{v}_i)]_{B'} = [a_{1,i}, a_{2,i}, \ldots, a_{m,i}]^T$. Hence the ij entry of $[T]_{B,B'}$ is $a_{i,j}$.

We often describe as follows.

$$T[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n] = [T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), \ldots, T(\boldsymbol{v}_n)] = [\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m][T]_{B,B'}.$$

It is because the ith column of the equation above reads

$$T(\boldsymbol{v}_i) = [\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m][T(\boldsymbol{v}_i)]_{B'} = a_{1,i}\boldsymbol{w}_1 + a_{2,i}\boldsymbol{w}_2 + \cdots + a_{m,i}\boldsymbol{w}_m.$$

Note that since $\boldsymbol{x} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n][\boldsymbol{x}]_B$,

$$T(\boldsymbol{x}) = T[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n][\boldsymbol{x}]_B = [\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m][T]_{B,B'}[\boldsymbol{x}]_B$$

and $[T(\boldsymbol{x})]_{B'} = [T]_{B,B'}[\boldsymbol{x}]_B$.

If $V = W$ and $B = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ and $B' = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ are bases and

$$\boldsymbol{u}_i = \sum_{j=1}^{n} p_{j,i}\boldsymbol{v}_j = p_{1,i}\boldsymbol{v}_1 + p_{2,i}\boldsymbol{v}_2 + \cdots + p_{n,i}\boldsymbol{v}_n,$$

then

$$[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n] = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]P, \text{ and } [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n] = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n]P^{-1}.$$

Hence

$$\begin{aligned} [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n][T]_{B'} &= T[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n] \\ &= T([\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]P) \\ &= (T[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n])P \quad \text{why?} \\ &= [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n][T]_B P \\ &= [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n]P^{-1}[T]_B P. \end{aligned}$$

Hence $[T]_{B'} = P^{-1}[T]_B P$.

9.3 Isomorphism Theorem

Definition 9.3 An *isomorphism* between V and W is a bijective linear transformation form V to W. When there is an isomorphism between V and W, we say V and W are isomorphic.

When $V = W$, isomorphisms are called *automorphisms*.

Proposition 9.5 *Let V and W be finite-dimensional vector space. Then V and W are isomorphic, i.e., there is a bijective linear transformation from V to W if and only if* $\dim V = \dim W$. *In particular, every real vector space of dimension n is isomorphic to $\boldsymbol{R}^n$.*

Proof. Let B be a basis of V. Then $T : V \to \boldsymbol{R}^n$ ($\boldsymbol{x} \mapsto [\boldsymbol{x}]_B$) is an isomorphism. Hence V is isomorphic to W if and only if $\boldsymbol{R}^n$ and $\boldsymbol{R}^m$ are isomorphic. If $n = m$, $\boldsymbol{R}^n$ and $\boldsymbol{R}^m$ are isomophic. Suppose $T : \boldsymbol{R}^n \to \boldsymbol{R}^m$ is an isomorphism. then $n = \mathrm{rank}(T) + \mathrm{nullity}(T) = \mathrm{rank}(T) = m$ by Proposition 8.7, as desired. ∎