# Vector Spaces and Linear Transformations

Beifang Chen

Fall 2006

## 1 Vector spaces

A **vector space** is a nonempty set $V$, whose objects are called **vectors**, equipped with two operations, called **addition** and **scalar multiplication**: For any two vectors $\boldsymbol{u}$, $\boldsymbol{v}$ in $V$ and a scalar $c$, there are unique vectors $\boldsymbol{u}+\boldsymbol{v}$ and $c\boldsymbol{u}$ in $V$ such that the following properties are satisfied.

1. $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$,
2. $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$,
3. There is a vector $\mathbf{0}$, called the **zero vector**, such that $\boldsymbol{u}+\mathbf{0}=\boldsymbol{u}$,
4. For any vector $\boldsymbol{u}$ there is a vector $-\boldsymbol{u}$ such that $\boldsymbol{u}+(-\boldsymbol{u})=\mathbf{0}$;
5. $c(\boldsymbol{u}+\boldsymbol{v})=c\boldsymbol{u}+c\boldsymbol{v}$,
6. $(c+d)\boldsymbol{u}=c\boldsymbol{u}+d\boldsymbol{u}$,
7. $c(d\boldsymbol{u})=(cd)\boldsymbol{u}$,
8. $1\boldsymbol{u}=\boldsymbol{u}$.

By definition of vector space it is easy to see that for any vector $\boldsymbol{u}$ and scalar $c$,

$$0\boldsymbol{u}=\mathbf{0},\quad c\mathbf{0}=\mathbf{0},\quad -\boldsymbol{u}=(-1)\boldsymbol{u}.$$

For instance,

$$\begin{aligned}
0\boldsymbol{u} &\overset{(3)}{=} 0\boldsymbol{u}+\mathbf{0} \overset{(4)}{=} 0\boldsymbol{u}+(0\boldsymbol{u}+(-0\boldsymbol{u})) \overset{(2)}{=} (0\boldsymbol{u}+0\boldsymbol{u})+(-0\boldsymbol{u}) \\
&\overset{(6)}{=} (0+0)\boldsymbol{u}+(-0\boldsymbol{u}) = 0\boldsymbol{u}+(-0\boldsymbol{u}) \overset{(4)}{=} \mathbf{0}; \\
c\mathbf{0} &= c(0\boldsymbol{u}) \overset{(7)}{=} (c0)\boldsymbol{u} = 0\boldsymbol{u} = \mathbf{0}; \\
-\boldsymbol{u} &= -\boldsymbol{u}+\mathbf{0} = -\boldsymbol{u}+(1-1)\boldsymbol{u} = -\boldsymbol{u}+\boldsymbol{u}+(-1)\boldsymbol{u} = \mathbf{0}+(-1)\boldsymbol{u} = (-1)\boldsymbol{u}.
\end{aligned}$$

**Example 1.1.** (a) The Euclidean space $\mathbb{R}^n$ is a vector space under the ordinary addition and scalar multiplication.

(b) The set $\mathbf{P}_n$ of all polynomials of degree less than or equal to $n$ is a vector space under the ordinary addition and scalar multiplication of polynomials.

(c) The set $\mathbf{M}(m,n)$ of all $m\times n$ matrices is a vector space under the ordinary addition and scalar multiplication of matrices.

(d) The set $C[a,b]$ of all continuous functions on the closed interval $[a,b]$ is a vector space under the ordinary addition and scalar multiplication of functions.

**Definition 1.1.** Let $V$ and $W$ be vector spaces, and $W\subseteq V$. If the addition and scalar multiplication in $W$ are the same as the addition and scalar multiplication in $V$, then $W$ is called a **subspace** of $V$.

If $H$ is a subspace of $V$, then $H$ is closed for the addition and scalar multiplication of $V$, i.e., for any $\boldsymbol{u}, \boldsymbol{v} \in H$ and scalar $c \in \mathbb{R}$, we have

$$\boldsymbol{u} + \boldsymbol{v} \in H, \quad c\boldsymbol{v} \in H.$$

For a nonempty set $S$ of a vector space $V$, to verify whether $S$ is a subspace of $V$, it is required to check (1) whether the addition and scalar multiplication are well defined in the given subset $S$, that is, whether they are closed under the addition and scalar multiplication of $V$; (2) whether the eight properties (1-8) are satisfied. However, the following theorem shows that we only need to check (1), that is, to check whether the addition and scalar multiplication are closed in the given subset $S$.

**Theorem 1.2.** *Let $H$ be a nonempty subset of a vector space $V$. Then $H$ is a subspace of $V$ if and only if $H$ is closed under addition and scalar multiplication, i.e.,*

(a) *For any vectors $\boldsymbol{u}, \boldsymbol{v} \in H$, we have $\boldsymbol{u} + \boldsymbol{v} \in H$,*

(b) *For any scalar $c$ and a vector $\boldsymbol{v} \in H$, we have $c\boldsymbol{v} \in H$.*

**Example 1.2.** (a) For a vector space $V$, the set $\{\mathbf{0}\}$ of the zero vector and the whole space $V$ are subspaces of $V$; they are called the **trivial subspaces** of $V$.

(b) For an $m \times n$ matrix $A$, the set of solutions of the linear system $A\boldsymbol{x} = \mathbf{0}$ is a subspace of $\mathbb{R}^n$. However, if $\boldsymbol{b} \neq \mathbf{0}$, the set of solutions of the system $A\boldsymbol{x} = \boldsymbol{b}$ is *not* a subspace of $\mathbb{R}^n$.

(c) For any vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ in $\mathbb{R}^n$, the span $\operatorname{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ is a subspace of $\mathbb{R}^n$.

(d) For any linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, the image

$$T(\mathbb{R}^n) = \{T(\boldsymbol{x}) : \boldsymbol{x} \in \mathbb{R}^n\}$$

of $T$ is a subspace of $\mathbb{R}^m$, and the inverse image

$$T^{-1}(\mathbf{0}) = \{\boldsymbol{x} \in \mathbb{R}^n : T(\boldsymbol{x}) = \mathbf{0}\}$$

is a subspace of $\mathbb{R}^n$.

# 2 Some special subspaces

**Lecture 15**

Let $A$ be an $m \times n$ matrix. The **null space** of $A$, denoted by $\operatorname{Nul} A$, is the space of solutions of the linear system $A\boldsymbol{x} = \mathbf{0}$, that is,

$$\operatorname{Nul} A = \{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} = \mathbf{0}\}.$$

The **column space** of $A$, denoted by $\operatorname{Col} A$, is the span of the column vectors of $A$, that is, if $A = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n]$, then

$$\operatorname{Col} A = \operatorname{Span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}.$$

The **row space** of $A$ is the span of the row vectors of $A$, and is denoted by $\operatorname{Row} A$.

Let $T : \mathbb{R}^n \to \mathbb{R}^m$, $T(\boldsymbol{x}) = A\boldsymbol{x}$, be a linear transformation. Then $\operatorname{Nul} A$ is the set of inverse images of $\mathbf{0}$ under $T$ and $\operatorname{Col} A$ is the image of $T$, that is,

$$\operatorname{Nul} A = T^{-1}(\mathbf{0}) \quad \text{and} \quad \operatorname{Col} A = T(\mathbb{R}^n).$$

## 3 Linear transformations

Let $V$ and $W$ be vector spaces. A function $T : V \to W$ is called a **linear transformation** if for any vectors $\boldsymbol{u}$, $\boldsymbol{v}$ in $V$ and scalar $c$,

(a) $T(\boldsymbol{u}+\boldsymbol{v}) = T(\boldsymbol{u}) + T(\boldsymbol{v})$,

(b) $T(c\boldsymbol{u}) = cT(\boldsymbol{u})$.

The inverse images $T^{-1}(\mathbf{0})$ of $\mathbf{0}$ is called the **kernel** of $T$ and $T(V)$ is called the **range** of $T$.

**Example 3.1.** (a) Let $A$ is an $m \times m$ matrix and $B$ an $n \times n$ matrix. The function

$$F : \mathbf{M}(m,n) \to \mathbf{M}(m,n), \quad F(X) = AXB$$

is a linear transformation. For instance, for $m = n = 2$, let

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}.$$

Then $F : \mathbf{M}(2,2) \to \mathbf{M}(2,2)$ is given by

$$F(X) = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 2x_1 + 2x_2 + 4x_3 + 4x_4 & x_1 + 3x_2 + 2x_3 + 6x_4 \\ 2x_1 + 2x_2 + 6x_3 + 6x_4 & x_1 + 3x_2 + 3x_3 + 9x_4 \end{bmatrix}.$$

(b) The function $D : \mathbf{P}_3 \to \mathbf{P}_2$, defined by

$$D\big(a_0 + a_1 t + a_2 t^2 + a_3 t^3\big) = a_1 + 2a_2 t + 3a_3 t^2,$$

is a linear transformation.

**Proposition 3.1.** *Let $T : V \to W$ be a linear transformation. Then $T^{-1}(\mathbf{0})$ is a subspace of $V$ and $T(V)$ is a subspace of $W$. Moreover,*

(a) *If $V_1$ is a subspace of $V$, then $T(V_1)$ is a subspace of $W$;*

(b) *If $W_1$ is a subspace of $W$, then $T^{-1}(W_1)$ is a subspace of $V$.*

*Proof.* By definition of subspaces. □

**Theorem 3.2.** *Let $T : V \to W$ be a linear transformation. Given vectors $v_1, v_2, \ldots, v_k$ in $V$.*

(a) *If $v_1, v_2, \ldots, v_k$ are linearly dependent, then $T(v_1), T(v_2), \ldots, T(v_k)$ are linearly dependent;*

(b) *If $T(v_1), T(v_2), \ldots, T(v_k)$ are linearly independent, then $v_1, v_2, \ldots, v_k$ are linearly independent.*

## 4 Independent sets and bases

**Definition 4.1.** Vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p$ of a vector space $V$ are called **linearly independent** if, whenever there are constants $c_1, c_2, \ldots, c_p$ such that

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_p\boldsymbol{v}_p = \mathbf{0},$$

we have

$$c_1 = c_2 = \cdots = c_p = 0.$$

The vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p$ are called **linearly dependent** if there exist constants $c_1, c_2, \ldots, c_p$, not all zero, such that

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_p\boldsymbol{v}_p = \mathbf{0}.$$

Any family of vectors that contains the zero vector $\mathbf{0}$ is linearly dependent. A single vector $\boldsymbol{v}$ is linearly independent if and only if $\boldsymbol{v} \neq \mathbf{0}$.

**Theorem 4.2.** *Vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ ($k \geq 2$) are linearly dependent if and only if one of the vectors is a linear combination of the others, i.e., there is one $i$ such that*

$$\boldsymbol{v}_i = a_1\boldsymbol{v}_1 + \cdots + a_{i-1}\boldsymbol{v}_{i-1} + a_{i+1}\boldsymbol{v}_{i+1} + \cdots + a_k\boldsymbol{v}_k.$$

*Proof.* Since the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ are linearly dependent, there are constants $c_1, c_2, \ldots, c_k$, not all zero, such that

$$c_1\boldsymbol{v} + c_2\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k = \mathbf{0}.$$

Let $c_i \neq 0$. Then

$$\boldsymbol{v}_i = \left(-\frac{c_1}{c_i}\right)\boldsymbol{v}_1 + \cdots + \left(-\frac{c_{i-1}}{c_i}\right)\boldsymbol{v}_{i-1} + \left(-\frac{c_{i+1}}{c_i}\right)\boldsymbol{v}_{i+1} + \cdots + \left(-\frac{c_k}{c_i}\right)\boldsymbol{v}_k.$$

□

**Note 1.** The condition $\boldsymbol{v}_1 \neq \mathbf{0}$ can not be omitted. For instance, the set $\{\mathbf{0}, \boldsymbol{v}_2\}$ ($\boldsymbol{v}_2 \neq \mathbf{0}$) is a dependent set, but $\boldsymbol{v}_2$ is not a linear combination of the zero vector $\boldsymbol{v}_1 = \mathbf{0}$.

**Theorem 4.3.** *Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ be a subset of independent vectors in a vector space $V$. If a vector $\boldsymbol{v}$ can be written in two linear combinations of $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$, say,*

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_k\boldsymbol{v}_k = d_1\boldsymbol{v}_1 + d_2\boldsymbol{v}_2 + \cdots + d_k\boldsymbol{v}_k,$$

*then*

$$c_1 = d_1, \quad c_2 = d_2, \quad \ldots, \quad c_k = d_k.$$

*Proof.* If $(c_1, c_2, \ldots, c_k) \neq (d_1, d_2, \ldots, d_k)$, then one of the entries in $(c_1 - d_1, c_2 - d_2, \ldots, c_k - d_k)$ is nonzero, and

$$(c_1 - d_1)\boldsymbol{v}_1 + (c_2 - d_2)\boldsymbol{v}_2 + \cdots + (c_k - d_k)\boldsymbol{v}_k = \mathbf{0}.$$

This means that the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ are linearly dependent. This is a contradiction. □

**Definition 4.4.** Let $H$ be a subspace of a vector space $V$. An ordered set $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ of vectors in $V$ is called a **basis** for $H$ if

(a) $\mathcal{B}$ is a linearly independent set, and

(b) $\mathcal{B}$ spans $H$, that is, $H = \operatorname{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$.

**Example 4.1.** (a) The set $\{1, t, t^2\}$ is basis of $\mathbf{P}_2$.
(b) The set $\{1, t+1, t^2+t\}$ is basis of $\mathbf{P}_2$.
(c) The set $\{1, t+1, t-1\}$ is not a basis of $\mathbf{P}_2$.

**Proposition 4.5 (Spanning Theorem).** *Let $H$ be a nonzero subspace of a vector space $V$ and $H = \operatorname{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$.*

(a) *If some $\boldsymbol{v}_k$ is a linear combination of the other vectors of $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$, then*

$$H = \operatorname{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_p\}.$$

(b) *If $H \neq \{0\}$, then some subsets of $S = \{v_1, v_2, \ldots, v_p\}$ are bases for $H$.*

*Proof.* It is clear that $\operatorname{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_p\}$ is contained in $H$. Write

$$\boldsymbol{v}_k = c_1\boldsymbol{v}_1 + \cdots + c_{k-1}\boldsymbol{v}_{k-1} + c_{k+1}\boldsymbol{v}_{k+1} + \cdots + c_p\boldsymbol{v}_p.$$

Then for any vector $\boldsymbol{v} = a_1\boldsymbol{v}_1 + \cdots + a_k\boldsymbol{v}_k + \cdots + a_p\boldsymbol{v}_p$ in $H$, we have

$$\begin{aligned}\boldsymbol{v} &= (a_1 + a_kc_1)\boldsymbol{v}_1 + \cdots + (a_{k-1} + a_kc_{k-1})\boldsymbol{v}_{k-1} \\ &\quad + (a_{k+1} + a_kc_{k+1})\boldsymbol{v}_{k+1} + \cdots + (a_p + a_kc_p)\boldsymbol{v}_p.\end{aligned}$$

This means that $\boldsymbol{v}$ is contained in $\operatorname{Span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{k-1}, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_p\}$. □

**Example 4.2.** The vector space $\mathbf{P}(t)$ of polynomials of degree $\le 2$ has a basis $\mathcal{B} = \{1, t, t^2\}$. The set $\mathcal{B}_1 = \{1, t+1, t^2\}$ is also a basis of $\mathbf{P}(t)$. However, $\{1, t+1, t-1\}$ is not a basis of $\mathbf{P}(t)$.

**Example 4.3.** The vector space $\mathbf{M}(2,2)$ of $2 \times 2$ matrices has a basis

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

The following set

$$\mathcal{C} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$$

is also a basis of $\mathbf{M}(2,2)$.

# 5 Bases of null and column spaces

**Example 5.1.** Consider the matrix

$$A = \begin{bmatrix} 1 & 4 & 0 & 2 & 4 \\ -1 & -4 & 1 & -3 & -2 \\ 2 & 8 & 1 & 3 & 10 \\ 1 & 4 & 1 & 1 & 6 \end{bmatrix} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4, \boldsymbol{a}_5].$$

Its reduced row echelon form is the matrix

$$B = \begin{bmatrix} 1 & 4 & 0 & 2 & 4 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = [\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3, \boldsymbol{b}_4, \boldsymbol{b}_5].$$

Since $A\boldsymbol{x} = \mathbf{0}$ is equivalent to $B\boldsymbol{x} = \mathbf{0}$, that is,

$$x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3 + x_4\boldsymbol{a}_4 + x_5\boldsymbol{a}_5 = \mathbf{0} \iff x_1\boldsymbol{b}_1 + x_2\boldsymbol{b}_2 + x_3\boldsymbol{b}_3 + x_4\boldsymbol{b}_4 + x_5\boldsymbol{b}_5 = \mathbf{0}.$$

This means that the linear relations among the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_4, \boldsymbol{a}_5$ are the same as the linear relations among the vectors $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3, \boldsymbol{b}_4, \boldsymbol{b}_5$. For instance,

$$\begin{array}{lcl} \boldsymbol{b}_2 = 4\boldsymbol{b}_1 & \longleftrightarrow & \boldsymbol{a}_2 = 4\boldsymbol{a}_1 \\ \boldsymbol{b}_4 = 2\boldsymbol{b}_1 - \boldsymbol{b}_3 & \longleftrightarrow & \boldsymbol{a}_4 = 2\boldsymbol{a}_1 - \boldsymbol{a}_3 \\ \boldsymbol{b}_5 = 4\boldsymbol{b}_1 + 2\boldsymbol{b}_3 & \longleftrightarrow & \boldsymbol{a}_4 = 4\boldsymbol{a}_1 + 2\boldsymbol{a}_3. \end{array}$$

This shows that row operations *do not* change the linear relations among the column vectors of a matrix.

**Note 2.** Let $A$ and $B$ be matrix such that $A \sim B$, that is, $A$ is equivalent to $B$. Then

$$\operatorname{Nul} A = \operatorname{Nul} B, \quad \operatorname{Row} A = \operatorname{Row} B, \quad \text{but} \quad \operatorname{Col} A \ne \operatorname{Col} B.$$

**Theorem 5.1 (Column Space Theorem).** *The column vectors of a matrix $A$ corresponding to its pivot positions form a basis of* $\operatorname{Col} A$.

*Proof.* Let $B = [\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n]$ denote the reduced row echelon form of $A = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n]$. Let $\boldsymbol{b}_{i_1}, \boldsymbol{b}_{i_2}, \ldots, \boldsymbol{b}_{i_k}$ be the column vectors of $B$ containing the pivot positions. It is clear that $\boldsymbol{b}_{i_1}, \boldsymbol{b}_{i_2}, \ldots, \boldsymbol{b}_{i_k}$ are linearly independent and every column vector of $B$ is a linear combination of the vectors $\boldsymbol{b}_{i_1}, \boldsymbol{b}_{i_2}, \ldots, \boldsymbol{b}_{i_k}$.

Let $\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, \ldots, \boldsymbol{a}_{i_k}$ be the corresponding column vectors of $A$. It suffices to prove that a linear relation for $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n$ is also a linear relation for $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$, and vice versa. Notice that a linear relation among the vectors $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n$ is just a solution of the system $B\boldsymbol{x} = \mathbf{0}$; and the systems $A\boldsymbol{x} = \mathbf{0}$ and $B\boldsymbol{x} = \mathbf{0}$ have the same solution set. Thus $\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, \ldots, \boldsymbol{a}_{i_k}$ are linearly independent and every column vector of $A$ is a linear combination of $\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, \ldots, \boldsymbol{a}_{i_k}$. So $\boldsymbol{a}_{i_1}, \boldsymbol{a}_{i_2}, \ldots, \boldsymbol{a}_{i_k}$ form a basis of $\operatorname{Col} A$. □

## 6 Coordinate systems

**Lecture 17**

**Theorem 6.1.** *Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis of a vector space $V$. Then for each vector $v$ in $V$, there exists a unique set of scalars $c_1, c_2, \ldots, c_n$ such that*

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n.$$

*Proof.* Trivial. □

Let $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis of a vector space $V$. Then every vector $\boldsymbol{v}$ of $V$ has a unique expression

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n.$$

The scalars $c_1, c_2, \ldots, c_n$ are called the **coordinates of $\boldsymbol{v}$ relative to the basis $\mathcal{B}$** (or **$\mathcal{B}$-coordinates of $\boldsymbol{v}$**); and the vector

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

is called the **coordinate vector of $\boldsymbol{v}$ relative to $\mathcal{B}$** (or the **$\mathcal{B}$-coordinate vector of $\boldsymbol{v}$**). We may write

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n = \big[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\big] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \mathcal{B}[\boldsymbol{v}]_{\mathcal{B}}.$$

**Example 6.1.** Any two linearly independent vectors of $\mathbb{R}^2$ form a basis for $\mathbb{R}^2$. For instance, the set

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

is basis of $\mathbb{R}^2$. The vector $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ has the $\mathcal{B}$-coordinate vector $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$. However, the coordinate vector of $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ is just itself under the standard basis

$$\mathcal{E} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}.$$

**Theorem 6.2.** *Let $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ be a basis of a subspace $H$ of $\mathbb{R}^n$. Let $P_{\mathcal{B}}$ be the matrix*

$$P_{\mathcal{B}} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p].$$

*Then for any vector $\boldsymbol{v}$ in $\mathbb{R}^n$,*

$$\boldsymbol{v} = P_{\mathcal{B}}[\boldsymbol{v}]_{\mathcal{B}}.$$

*The matrix $P_{\mathcal{B}}$, which transfers the $\mathcal{B}$-coordinate vector $[\boldsymbol{v}]_{\mathcal{B}}$ of $\boldsymbol{v}$ to its standard coordinate vector $\boldsymbol{v} = [\boldsymbol{v}]_{\mathcal{E}}$, is called the* **change-of-coordinate matrix from $\mathcal{B}$ to $\mathcal{E}$**.

*Proof.* Let $\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n$. Then

$$\boldsymbol{v} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_p \end{bmatrix} = P_{\mathcal{B}}[\boldsymbol{v}]_{\mathcal{B}}.$$

□

**Theorem 6.3.** *Let $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis of a vector space $V$. Then the coordinate transformation,*

$$V \to \mathbb{R}^n, \quad \boldsymbol{v} \mapsto [\boldsymbol{v}]_{\mathcal{B}},$$

*is linear, one-to-one, and onto.*

*Proof.* For vectors $\boldsymbol{v}, \boldsymbol{w}$ of $V$ and scalar $a$, if

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n,$$

$$\boldsymbol{w} = d_1\boldsymbol{v}_1 + d_2\boldsymbol{w}_2 + \cdots + d_n\boldsymbol{v}_n,$$

then

$$\boldsymbol{v} + \boldsymbol{w} = (c_1 + d_1)\boldsymbol{v}_1 + (c_2 + d_2)\boldsymbol{v}_2 + \cdots + (c_n + d_n)\boldsymbol{v}_n,$$

$$a\boldsymbol{v} = a(c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n) = (ac_1)\boldsymbol{v}_1 + (ac_2)\boldsymbol{v}_2 + \cdots + (ac_n)\boldsymbol{v}_n.$$

Thus

$$[\boldsymbol{v} + \boldsymbol{w}]_{\mathcal{B}} = \begin{bmatrix} c_1 + d_1 \\ c_2 + d_2 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix},$$

$$[c\boldsymbol{v}]_{\mathcal{B}} = \begin{bmatrix} cc_1 \\ \vdots \\ cc_n \end{bmatrix} = c \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = c[\boldsymbol{v}]_{\mathcal{B}}.$$

So the coordinate transformation is a linear transformation.

Now for any vector

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \quad \text{in} \quad \mathbb{R}^n,$$

consider the vector $\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n$ in $V$. The coordinate vector of $\boldsymbol{v}$ relative to $\mathcal{B}$ is

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}.$$

Thus the transformation is onto. The linear independence of $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ implies that the transformation is also one-to-one. □

A one-to-one and onto linear transformation from a vector space $V$ to a vector space $W$ is called an **isomorphism**.

**Example 6.2.** The vector space $\mathbf{P}_3$ of polynomials of degree at most 3 in variable $t$ is isomorphic to the vector space $\mathbb{R}^4$, and $\{1, t, t^2, t^3\}$ is a basis of $\mathbf{P}_3$.

*Proof.* The map $F : \mathbf{P}_3 \to \mathbb{R}^4$, defined by

$$F[p(t)] = F\left(c_0 + c_1 t + c_2 t^2 + c_3 t^3\right) = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix},$$

is a one-to-one linear transformation from $\mathbf{P}_3$ onto $\mathbb{R}^4$. □

**Example 6.3.** The vector space $\mathbf{M}(2,2)$ of $2\times 2$ matrices is isomorphic to $\mathbb{R}^4$, and the matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

form a basis of $\mathbf{M}(2,2)$. In fact, the map $F : \mathbf{M}(2,2) \to \mathbb{R}^4$, defined by

$$F(M) = F\left(\begin{bmatrix} a_1 & a_2 \\ a_3 & a_4 \end{bmatrix}\right) = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix},$$

is a one-to-one linear transformation from $\mathbf{M}(2,2)$ onto $\mathbb{R}^4$.

**Theorem 6.4.** *Let $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis of a vector space $V$. Then any set of $V$ consisting more than $n$ vectors are linearly dependent.*

*Proof.* Let $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_p\}$ be a set of vectors with $p > n$. Since any set of more than $n$ vectors of $\mathbb{R}^n$ is linearly dependent, the vectors $[\boldsymbol{u}_1]_{\mathcal{B}}, [\boldsymbol{u}_2]_{\mathcal{B}}, \ldots, [\boldsymbol{u}_p]_{\mathcal{B}}$ of $\mathbb{R}^n$ must be linearly dependent. Then there exist constants $c_1, c_2, \ldots, c_p$, not all zero, such that

$$c_1[\boldsymbol{u}_1]_{\mathcal{B}} + c_2[\boldsymbol{u}_2]_{\mathcal{B}} + \cdots + c_p[\boldsymbol{u}_p]_{\mathcal{B}} = \mathbf{0}.$$

Thus

$$[c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_p\boldsymbol{u}_p]_{\mathcal{B}} = c_1[\boldsymbol{u}_1]_{\mathcal{B}} + c_2[\boldsymbol{u}_2]_{\mathcal{B}} + \cdots + c_p[\boldsymbol{u}_p]_{\mathcal{B}} = \mathbf{0} = [\mathbf{0}]_{\mathcal{B}}.$$

Note that the coordinate transformation is one-to-one. It follows that

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_p\boldsymbol{u}_p = \mathbf{0}.$$

This means that the vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_p$ are linearly dependent by definition. □

**Theorem 6.5.** *If $\mathcal{B}_1 = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ and $\mathcal{B}_2 = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ are bases of a vector space $V$, then $n = p$.*

*Proof.* Suppose $n < p$. By Theorem 6.4, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_p\}$ is linearly dependent, contrary to the properties for a basis. Thus $n \geq p$. A similar argument shows that $n \leq p$. Hence $n = p$. □

# 7 Dimensions of vector spaces

**Lecture 18**

A vector space $V$ is said to be **finite dimensional** if it can be spanned by a set of finite number of vectors. The dimension of $V$, denoted by $\dim V$, is the number of vectors of a basis of $V$. The dimension of the zero vector space $\{\mathbf{0}\}$ is zero. If $V$ cannot be spanned by any finite set of vectors, then $V$ is said to be **infinite dimensional**.

**Theorem 7.1.** *Let $H$ be a subspace of a finite dimensional vector space $V$. Then any linearly independent subset of $H$ can be expanded to a basis of $H$. Moreover, $H$ is finite dimensional and* $\dim H \leq \dim V$.

*Proof.* Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ be a set of linearly independent vectors of $H$. If $\operatorname{Span} S = H$, then $S$ is a basis of $H$ by definition. Otherwise, there exists a vector $\boldsymbol{v}_{k+1}$ in $H$ such that $\boldsymbol{v}_{k+1}$ is not in $\operatorname{Span} S$. Then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}\}$ is a linearly independent set of $H$. Now set $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{k+1}\}$. If $\operatorname{Span} S = H$, then $S$ is a basis of $H$. Otherwise, continue to add one vector of $H - \operatorname{Span} S$ to $S$ in this way until $\operatorname{Span} S = H$. Since $H$ is of finite dimensional, the extension ends in finite number of steps. □

**Theorem 7.2 (Basis Theorem).** *Given a set $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ of $n$ vectors of an $n$-dimensional vector space $V$.*

(a) *If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is linearly independent, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a basis of $V$.*

(b) *If* $\operatorname{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} = V$, *then* $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ *is a basis of* $V$.

*Proof.* (a) By Theorem 7.1, $S$ can be extended to a basis of $V$. Since $S$ has $n$ vectors and all bases have the same number of vectors. It follows that no vectors were added to $S$ to be extended a basis of $V$. Hence $S$ itself must be a basis.

(b) We need to show that $S$ is linearly independent. Note that if $S$ is not a basis, then $S$ is linearly dependent. Thus $S$ contains a linearly independent proper subset $S'$ such that $\operatorname{Span} S' = V$. So $S'$ is a basis of $V$; therefore $\#(S') \geq n$, contradict to $\#(S') < n$. □

# 8 Rank

**Theorem 8.1.** *For any rectangular matrix* $A$,

$$\dim \operatorname{Row} A = \dim \operatorname{Col} A = \#(\text{pivot positions of } A).$$

**Definition 8.2.** The **rank** of a rectangular matrix $A$ is the number pivot positions of $A$, that is, the dimension of the row space and the column space of $A$. For a linear transformation $T : V \to W$, the **rank** of $T$ is the dimension of the subspace $T(V)$.

**Theorem 8.3 (Rank Theorem).** *For any* $m \times n$ *matrix* $A$,

$$\operatorname{rank} A + \dim \operatorname{Nul} A = n.$$

*Proof.* The rank of $A$ is the number of pivot positions of $A$ and the dimension of the null space of $A$ is the number of free variables of the system $A\boldsymbol{x} = \boldsymbol{0}$. It is clear that

$$\#(\text{pivot positions}) + \#(\text{free variables}) = n.$$

□

**Theorem 8.4.** *Let* $A$ *be an* $n \times n$ *invertible matrix. Then*

$$\dim \operatorname{Row} A = \dim \operatorname{Col} A = \operatorname{rank} A = n,$$

$$\dim \operatorname{Nul} A = 0.$$

*Proof.* The invertibility of $A$ implies that the number of pivot positions of $A$ is $n$. So $\operatorname{rank} A = n$ and $\dim \operatorname{Nul} A = 0$. □

# 9 Matrices of linear transformations

**Definition 9.1.** Let $T : V \to W$ be a linear transformation from a vector space $V$ with basis $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ to a vector space $W$ basis $\mathcal{C} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$. Let

$$\begin{cases} T(\boldsymbol{v}_1) &= a_{11}\boldsymbol{w}_1 + a_{21}\boldsymbol{w}_2 + \cdots + a_{m1}\boldsymbol{w}_m \\ T(\boldsymbol{v}_2) &= a_{12}\boldsymbol{w}_1 + a_{22}\boldsymbol{w}_2 + \cdots + a_{m2}\boldsymbol{w}_m \\ & \vdots \\ T(\boldsymbol{v}_n) &= a_{1n}\boldsymbol{w}_1 + a_{2n}\boldsymbol{w}_2 + \cdots + a_{mn}\boldsymbol{w}_m \end{cases}$$

writing more compactly,

$$\big[T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), \ldots, T(\boldsymbol{v}_n)\big] = \big[\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\big] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \big[\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\big] A.$$

The $m \times n$ matrix $A$ is called the **matrix of $T$ relative to the basis $\mathcal{B}$ of $V$ and the basis $\mathcal{C}$ of $W$.** Alternatively, the matrix $A$ can be defined as

$$A = \Big[[T(\boldsymbol{v}_1)]_{\mathcal{C}},\ [T(\boldsymbol{v}_2)]_{\mathcal{C}},\ \ldots,\ [T(\boldsymbol{v}_n)]_{\mathcal{C}}\Big].$$

Consider the isomorphisms

$$\begin{aligned} T_1 : V \longrightarrow \mathbb{R}^n \quad &\text{by} \quad T_1(\boldsymbol{v}) = [\boldsymbol{v}]_{\mathcal{B}}, \\ T_2 : W \longrightarrow \mathbb{R}^m \quad &\text{by} \quad T_2(\boldsymbol{w}) = [\boldsymbol{w}]_{\mathcal{C}}. \end{aligned} \tag{9.1}$$

The composition

$$T_2 \circ T \circ T_1^{-1} : \mathbb{R}^n \xrightarrow{T_1^{-1}} V \xrightarrow{T} W \xrightarrow{T_2} \mathbb{R}^m$$

is a linear transformation. For any vector $\boldsymbol{x}$ of $\mathbb{R}^n$, let $\boldsymbol{v}$ be the vector of $V$ whose coordinate vector is $\boldsymbol{x}$, i.e., $[\boldsymbol{v}]_{\mathcal{B}} = \boldsymbol{x}$, or

$$\boldsymbol{v} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n = \big[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\big]\boldsymbol{x}.$$

Then

$$\begin{aligned} T(\boldsymbol{v}) &= T(x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n) \\ &= x_1T(\boldsymbol{v}_1) + x_2T(\boldsymbol{v}_2) + \cdots + x_nT(\boldsymbol{v}_n) \\ &= \big[T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), \ldots, T(\boldsymbol{v}_n)\big]\boldsymbol{x} \\ &= \big[\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\big]A\boldsymbol{x}. \end{aligned}$$

The coordinate vector of $T(\boldsymbol{v})$ relative to the basis $\mathcal{C}$ is $A\boldsymbol{x}$. Thus

$$\big(T_2 \circ T \circ T_1^{-1}\big)(\boldsymbol{x}) = T_2\Big(T\big(T_1^{-1}(\boldsymbol{x})\big)\Big) = T_2\big(T(\boldsymbol{v})\big) = \big[T(\boldsymbol{v})\big]_{\mathcal{C}} = A\boldsymbol{x}.$$

This means that $A$ is the standard matrix of the linear transformation $T_2 \circ T \circ T_1^{-1} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$.

In particular, for a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, $T(\boldsymbol{x}) = A\boldsymbol{x}$. Let $\mathcal{E}_n = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ be the standard basis of $\mathbb{R}^n$, and let $\mathcal{E}_m = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m\}$ be the standard basis of $\mathbb{R}^m$. Then

$$\big[T(\boldsymbol{e}_1), T(\boldsymbol{e}_2), \ldots, T(\boldsymbol{e}_n)\big] = \big[\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m\big]A.$$

**Example 9.1.** Let $T : \mathbf{P}_3 \to \mathbf{P}_2$ be a linear transformation defined by

$$T\big(a_0 + a_1t + a_2t^2 + a_3t^3\big) = (a_0 + a_3) + (a_1 + a_2)t + (a_0 + a_1 + a_2 + a_3)t^2.$$

It is clear that $\mathbf{P}_3$ has a basis

$$\mathcal{B} = \big\{1,\ t,\ t(t+1),\ t(t+1)(t+2)\big\}$$

and $\mathbf{P}_2$ has a basis

$$\mathcal{C} = \big\{1,\ t,\ t(t-1)\big\}$$

The matrix of $T$ relative to the bases $\mathcal{B}$ and $\mathcal{C}$ can be found as follows:

$$\left\{\begin{array}{rcrcr} T(1) &=& 1 + t^2 &=& 1 + t + t(t-1) \\ T(t) &=& t + t^2 &=& 2t + t(t-1) \\ T\big((t(t+1)\big) &=& 2t + 2t^2 &=& 4t + 2t(t-1) \\ T\big(t(t+1)(t+2)\big) &=& 1 + 5t + 6t^2 &=& 1 + 11t + 6t(t-1) \end{array}\right.$$

which is equivalent to

$$\Big[T(1),\, T(t),\, T\big(t(t+1)\big),\, T\big(t(t+1)(t+2)\big)\Big] = \big[1,\, t,\, t(t-1)\big] \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 2 & 4 & 11 \\ 1 & 1 & 2 & 6 \end{bmatrix}.$$

Thus the matrix of $T$ relative to the bases $\mathcal{B}$ and $\mathcal{C}$ is

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 2 & 4 & 11 \\ 1 & 1 & 2 & 6 \end{bmatrix}.$$

**Theorem 9.2.** *Let $V$ be an $l$-dimensional subspace of $\mathbb{R}^n$ with a basis $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_l\}$, and let $W$ be a $k$-dimensional subspace of $\mathbb{R}^m$ with a basis $\mathcal{C} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}$. Let $T : V \to W$ be a linear transformation. Then the matrix $A$ of $T$ relative to the bases $\mathcal{B}$ and $\mathcal{C}$ is given by*

$$\Big[\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k \,\big|\, T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), \ldots, T(\boldsymbol{v}_l)\Big] \sim \left[\begin{array}{c|c} I_k & A \\ \hline 0 & 0 \end{array}\right].$$

**Corollary 9.3.** *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Let $\mathcal{B} = \big\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\big\}$ be a basis of $\mathbb{R}^n$, and let $\mathcal{C} = \big\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\big\}$ be a basis of $\mathbb{R}^m$. If*

$$\Big[\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m \,\big|\, T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), \ldots, T(\boldsymbol{v}_n)\Big] \sim \Big[I_m \,\big|\, A\Big],$$

*then $A$ is the matrix of $T$ relative to the bases $\mathcal{B}$ and $\mathcal{C}$.*

**Example 9.2.** Let $V$ be the subspace of $\mathbb{R}^4$ defined by the linear equation $x_1 + x_2 + x_3 + x_4 = 0$, and let $W$ be the subspace of $\mathbb{R}^3$ defined by $y_1 + 2y_2 + y_3 = 0$. Let $T : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ be defined by

$$T(x_1, x_2, x_3, x_4) = \big(3x_1 + x_2, x_2 + x_3, x_3 + 3x_4\big).$$

(a) Show that $T$ is a linear transformation from $V$ to $W$.

(b) Verify that

$$\mathcal{B} = \left\{ \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \ \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \ \boldsymbol{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix} \right\}$$

is basis of $V$, and

$$\mathcal{C} = \left\{ \boldsymbol{w}_1 = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}, \ \boldsymbol{w}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

is basis of $W$.

(c) Find the matrix of $T$ relative to the basis $\mathcal{B}$ of $V$ and the basis $\mathcal{C}$ of $W$.

*Solution.* (a) Let $[x_1, x_2, x_3, x_4]^T \in V$, i.e., $x_1 + x_2 + x_3 + x_4 = 0$. We have

$$(2x_1 + x_2) + 2(x_2 + x_3) + (x_3 + 2x_3) = 3(x_1 + x_2 + x_3 + x_4) = 0.$$

So $T$ is a well-defined linear transformation from $V$ to $W$. (b) Trivial. (c) Consider the matrix

$$\big[\boldsymbol{w}_1, \boldsymbol{w}_2 \,|\, T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), T(\boldsymbol{v}_3)\big] = \left[\begin{array}{rr|rrr} 0 & 1 & 2 & 3 & 3 \\ 1 & -1 & -1 & -1 & 0 \\ -2 & 1 & 0 & -1 & -3 \end{array}\right] \sim$$

$$\left[\begin{array}{rr|rrr} 1 & -1 & -1 & -1 & 0 \\ 0 & 1 & 2 & 3 & 3 \\ -2 & 1 & 0 & -1 & -3 \end{array}\right] \sim \left[\begin{array}{rr|rrr} 1 & -1 & -1 & -1 & 0 \\ 0 & 1 & 2 & 3 & 3 \\ 0 & -1 & -2 & -3 & -3 \end{array}\right] \sim \left[\begin{array}{rr|rrr} 1 & 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 & 3 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

The matrix of $T$ relative to the basis $\mathcal{B}$ and the basis $\mathcal{C}$ is

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 3 \end{bmatrix}.$$

# 10 Matrices of linear operators

Let $V$ be an $n$-dimensional vector space with a basis $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$. A linear transformation $T : V \to V$ is called a **linear operator** on $V$. The matrix

$$A = \Big[ \big[T(\boldsymbol{v}_1)\big]_{\mathcal{B}}, \big[T(\boldsymbol{v}_2)\big]_{\mathcal{B}}, \ldots, \big[T(\boldsymbol{v}_n)\big]_{\mathcal{B}} \Big]$$

is called the **matrix of $T$ relative to the basis $\mathcal{B}$**.

**Theorem 10.1.** *Let $T : V \to V$ be a linear transformation from a finite dimensional vector space $V$ to itself. Let $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ and $\mathcal{B}' = \{\boldsymbol{v}'_1, \boldsymbol{v}'_2, \ldots, \boldsymbol{v}'_n\}$ be bases of $V$ with the transition matrix $P$, that is,*

$$\big[\boldsymbol{v}'_1, \boldsymbol{v}'_2, \ldots, \boldsymbol{v}'_n\big] = \big[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\big] P.$$

*Let $A$ be the matrix of $T$ relative to the basis $\mathcal{B}$, and let $A'$ be the matrix of $T$ relative to the basis $\mathcal{B}'$. Then*

$$A' = P^{-1}AP.$$

**Corollary 10.2.** *Let $T : \mathbb{R}^n \to \mathbb{R}^n$, $T\boldsymbol{x} = A\boldsymbol{x}$, be a linear transformation. Let $\mathcal{E} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ be the standard basis of $\mathbb{R}^n$, and let $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be another basis. Then*

$$\big[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\big] = \big[\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\big] P,$$

*where $P = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n]$, and the matrix of $T$ relative to the basis $\mathcal{B}$ is*

$$P^{-1}AP.$$

**Example 10.1.** Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation given by

$$T(\boldsymbol{x}) = \begin{bmatrix} 7 & -10 \\ 5 & -8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Find the matrix of $T$ relative to the basis

$$\mathcal{B} = \left\{ \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \boldsymbol{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}.$$

*Solution.* Since

$$\begin{aligned} T(\boldsymbol{v}_1) &= \begin{bmatrix} 7 & -10 \\ 5 & -8 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ -3 \end{bmatrix} = -3\boldsymbol{v}_1 = [\boldsymbol{v}_1, \boldsymbol{v}_2] \begin{bmatrix} -3 \\ 0 \end{bmatrix}, \\ T(\boldsymbol{v}_2) &= \begin{bmatrix} 1 & -10 \\ 5 & -8 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 2\boldsymbol{v}_2 = [\boldsymbol{v}_1, \boldsymbol{v}_2] \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \end{aligned}$$

then

$$[T(\boldsymbol{v}_1), T(\boldsymbol{v}_2)] = [\boldsymbol{v}_1, \boldsymbol{v}_2] \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}.$$

The matrix of $T$ relative to the basis $\mathcal{B}$ is $\begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}$.

Let $\mathcal{E} = \{\boldsymbol{e}_1, \boldsymbol{e}_2\}$. Then $\big[\boldsymbol{v}_1, \boldsymbol{v}_2\big] = \big[\boldsymbol{e}_1, \boldsymbol{e}_2\big] P$, where $P = \big[\boldsymbol{v}_1, \boldsymbol{v}_2\big] = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$. One verifies

$$\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 7 & -10 \\ 5 & -8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}.$$

We say that the matrix the matrix $\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ diagonalizes the matrix $\begin{bmatrix} 7 & -10 \\ 5 & -8 \end{bmatrix}$.

**Definition 10.3.** An $n \times n$ matrix $A$ is said to be **similar** to an $n \times n$ matrix $B$ if there is an invertible matrix $P$ such that

$$P^{-1}AP = B.$$

**Theorem 10.4.** *For a finite dimensional vector space $V$ and a linear transformation $T : V \to V$, the matrices of $T$ relative to various bases are similar. In other words, the matrices of the same linear transformation from a vector space to itself under different bases are similar.*

## 11 Change of basis

Let $\mathcal{B} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ and $\mathcal{C} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be bases of a vector space $V$. Then

$$\begin{aligned} \boldsymbol{v}_1 &= p_{11}\boldsymbol{u}_1 + p_{21}\boldsymbol{u}_2 + \cdots + p_{n1}\boldsymbol{u}_n, \\ \boldsymbol{v}_2 &= p_{12}\boldsymbol{u}_1 + p_{22}\boldsymbol{u}_2 + \cdots + p_{n2}\boldsymbol{u}_n, \\ &\vdots \\ \boldsymbol{v}_n &= p_{1n}\boldsymbol{u}_1 + p_{2n}\boldsymbol{u}_2 + \cdots + p_{nn}\boldsymbol{u}_n. \end{aligned}$$

We may write this more compactly as

$$\big[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\big] = \big[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\big] \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} = \big[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\big] P.$$

The matrix $P$ is called the **transition matrix** from the basis $\mathcal{B}$ to the basis $\mathcal{C}$. Clearly, the transition matrix from the basis $\mathcal{C}$ to the basis $\mathcal{B}$ is the inverse matrix $P^{-1}$.

Let $\mathcal{B} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ and $\mathcal{C} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\}$ be bases for a $k$-dimensional subspace $V$ of $\mathbb{R}^n$. Then there is a $k \times k$ invertible matrix $P$ such that

$$\big[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\big] = \big[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\big] P.$$

The matrix $P$ is called the **transition matrix** from the basis $\mathcal{B}$ to the basis $\mathcal{C}$. For a vector $\boldsymbol{v}$ in $V$, let $\boldsymbol{x}$ and $\boldsymbol{y}$ be the coordinate vectors of $\boldsymbol{v}$ relative to the bases $\mathcal{B}$ and $\mathcal{C}$, respectively, i.e.,

$$\boldsymbol{v} = \big[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\big]\boldsymbol{x} = \big[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\big]\boldsymbol{y}.$$

Then

$$\big[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\big]\boldsymbol{x} = \big[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\big] P\boldsymbol{y}.$$

It follows that

$$\boldsymbol{x} = P\boldsymbol{y} \quad \text{or} \quad \boldsymbol{y} = P^{-1}\boldsymbol{x}.$$

The matrix $P^{-1}$ is also called the **change-of-coordinate matrix** from the basis $\mathcal{B}$ to the basis $\mathcal{C}$.

To find the invertible matrix $P$, we set

$$P = [\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_k], \quad B = [\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k], \quad C = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k]$$

Then

$$B\boldsymbol{p}_1 = \boldsymbol{v}_1, \quad B\boldsymbol{p}_2 = \boldsymbol{v}_2, \quad \ldots, \quad B\boldsymbol{p}_k = \boldsymbol{v}_k.$$

This means that $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_k$ are solutions of the linear systems

$$B\boldsymbol{x} = \boldsymbol{v}_1, \quad B\boldsymbol{x} = \boldsymbol{v}_2, \quad \ldots, \quad B\boldsymbol{x} = \boldsymbol{v}_k$$

respectively. The linear systems can be solved simultaneously as follows:

$$\big[\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k \,|\, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k\big] \sim \left[\begin{array}{c|c} I_k & P \\ \hline 0 & 0 \end{array}\right].$$

**Example 11.1.** Let $V$ be a subspace of $\mathbb{R}^5$ with two bases

$$\mathcal{B} = \left\{ \begin{bmatrix} 6 \\ 6 \\ 2 \\ 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 5 \\ 7 \\ 3 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ -1 \\ 2 \\ 3 \end{bmatrix} \right\}, \quad \mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \\ 2 \end{bmatrix} \right\}.$$

Let $\boldsymbol{v}$ be a vector of $V$ whose coordinate vector relative to the basis $\mathcal{B}$ is $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Find the coordinate vector of $\boldsymbol{v}$ relative to the basis $\mathcal{C}$.

*Solution.* Performing the row operations to the $n$-by-$n$ matrix $[A|B]$, we have

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 6 & 5 & 3 \\ 2 & 1 & 2 & 6 & 7 & 0 \\ 1 & 0 & 1 & 2 & 3 & -1 \\ 2 & 1 & 0 & 4 & 3 & 2 \\ 0 & 2 & 1 & 5 & 4 & 3 \end{array}\right] \quad \begin{array}{c} R_2 - 2R_1 \\ R_3 - R_1 \\ \to \\ R_4 - 2R_1 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 6 & 5 & 3 \\ 0 & -3 & 0 & -6 & -3 & -6 \\ 0 & -2 & 0 & -4 & -2 & -4 \\ 0 & -3 & -2 & -8 & -7 & -4 \\ 0 & 2 & 1 & 5 & 4 & 3 \end{array}\right] \quad \begin{array}{c} R_2/(-3) \\ \to \\ R_3 \leftrightarrow R_5 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 6 & 5 & 3 \\ 0 & 1 & 0 & 2 & 1 & 2 \\ 0 & 2 & 1 & 5 & 4 & 3 \\ 0 & -3 & -2 & -8 & -7 & -4 \\ 0 & -2 & 0 & -4 & -2 & -4 \end{array}\right] \quad \begin{array}{c} R_3 - 2R_2 \\ R_4 + 3R_2 \\ \to \\ R_5 + 2R_2 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 6 & 5 & 3 \\ 0 & 1 & 0 & 2 & 1 & 2 \\ 0 & 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & -2 & -2 & -4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \quad \begin{array}{c} R_4 + 2R_3 \\ R_1 - R_3 \\ \to \\ R_1 - 2R_2 \end{array}$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 2 & 1 & 2 \\ 0 & 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Then $B = CP$, where

$$P = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 1 & 2 & -1 \end{bmatrix}$$

So the coordinate vector of $\boldsymbol{v}$ relative to $\mathcal{C}$ is

$$\begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 1 & 2 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}.$$

**Example 11.2.** Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by

$$T(\boldsymbol{x}) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad \boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Consider two bases

$$\mathcal{B} = \left\{ \boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \boldsymbol{u}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{u}_3 = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix} \right\},$$

$$\mathcal{C} = \left\{ \boldsymbol{v}_1 = \begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix},\ \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix},\ \boldsymbol{v}_3 = \begin{bmatrix} 3 \\ -1 \\ 6 \end{bmatrix} \right\}.$$

Then

$$\big[\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\big] = \big[\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\big] \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 1 & 1 & 2 \end{bmatrix}.$$

**Theorem 11.1.** *Let $\mathcal{B} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ and $\mathcal{B}' = \{\boldsymbol{v}'_1, \boldsymbol{v}'_2, \ldots, \boldsymbol{v}'_n\}$ be bases of $V$ with the transition matrix $P$, that is,*

$$\big[\boldsymbol{v}'_1, \boldsymbol{v}'_2, \ldots, \boldsymbol{v}'_n\big] = \big[\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\big] P.$$

*Let $\mathcal{C} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$ and $\mathcal{C}' = \{\boldsymbol{w}'_1, \boldsymbol{w}'_2, \ldots, \boldsymbol{w}'_m\}$ be bases of $W$ with the connection matrix $Q$, that is,*

$$\big[\boldsymbol{w}'_1, \boldsymbol{w}'_2, \ldots, \boldsymbol{w}'_m\big] = \big[\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\big] Q.$$

*Let $T : V \to W$ be a linear transformation with the matrix $A$ relative to the bases $\mathcal{B}$ and $\mathcal{C}$, then the matrix of $T$ relative to the bases $\mathcal{B}'$ and $\mathcal{C}'$ is given by*

$$Q^{-1}AP.$$

*Proof.* Let $A'$ denote the matrix of $T$ relative to the bases $\mathcal{B}'$ and $\mathcal{C}$. Then

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

On the other hand, since

$$\begin{aligned} \boldsymbol{v}'_1 &= p_{11}\boldsymbol{v}_1 + p_{21}\boldsymbol{v}_2 + \cdots + p_{n1}\boldsymbol{v}_n \\ \boldsymbol{v}'_2 &= p_{12}\boldsymbol{v}_1 + p_{22}\boldsymbol{v}_2 + \cdots + p_{n2}\boldsymbol{v}_n \\ &\vdots \\ \boldsymbol{v}'_n &= p_{1n}\boldsymbol{v}_1 + p_{2n}\boldsymbol{v}_2 + \cdots + p_{nn}\boldsymbol{v}_n \end{aligned}$$

then

$$\begin{aligned} T(\boldsymbol{v}'_1) &= p_{11}T(\boldsymbol{v}_1) + p_{21}T(\boldsymbol{v}_2) + \cdots + p_{n1}T(\boldsymbol{v}_n) \\ T(\boldsymbol{v}'_2) &= p_{12}T(\boldsymbol{v}_1) + p_{22}T(\boldsymbol{v}_2) + \cdots + p_{n2}T(\boldsymbol{v}_n) \\ &\vdots \\ T(\boldsymbol{v}'_n) &= p_{1n}T(\boldsymbol{v}_1) + p_{2n}T(\boldsymbol{v}_2) + \cdots + p_{nn}T(\boldsymbol{v}_n) \end{aligned}$$

Thus

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

Therefore $QA' = AP$, that is,

$$A' = Q^{-1}AP.$$

□

**Example 11.3.** Let $T : \mathbf{P}_3 \to \mathbf{P}_2$ be a linear transformation defined by

$$T(a_0 + a_1t + a_2t^2 + a_3t^3) = (a_0 + a_3) + (a_1 + a_2)t + (a_0 + a_1 + a_2 + a_3)t^2.$$

Then the matrix of $T$ relative to the bases

$$\mathcal{B} = \{1,\ t,\ t(t+1),\ t(t+1)(t+2)\} \quad \text{and} \quad \mathcal{C} = \{1,\ t,\ t(t-1)\}$$

is the matrix

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 2 & 4 & 11 \\ 1 & 1 & 2 & 6 \end{bmatrix}.$$

Given new bases

$$\mathcal{B}' = \{1,\ t,\ t(t-1),\ t(t-1)(t-2)\} \quad \text{and} \quad \mathcal{C}' = \{1,\ t,\ t(t+1)\}.$$

Since

$$\begin{cases} T(1) &=& 1+t^2 &=& 1-t+t(t+1) \\ T(t) &=& t+t^2 &=& t(t+1) \\ T\big(t(t-1)\big) &=& 0 &=& 0 \\ T\big(t(t-1)(t-2)\big) &=& 1-t &=& 1-t \end{cases},$$

the matrix of $T$ relative to the bases $\mathcal{B}'$ and $\mathcal{C}'$ is

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 0 & 0 & -1 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

Note that

$$\big[T(1),\ T(t),\ T\big(t(t+1)\big),\ T\big(t(t+1)(t+2)\big)\big] = [1,\ t,\ t(t-1)] \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 2 & 4 & 11 \\ 1 & 1 & 2 & 6 \end{bmatrix},$$

$$[1,\ t,\ t(t-1),\ t(t-1)(t-2)] = [1,\ t,\ t(t+1),\ t(t+1)(t+2)] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 6 \\ 0 & 0 & 1 & -6 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$[1,\ t,\ t(t+1)] = [1,\ t,\ t(t-1)] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

One verifies that

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ -1 & 0 & 0 & -1 \\ 1 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 2 & 4 & 11 \\ 1 & 1 & 2 & 6 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 6 \\ 0 & 0 & 1 & -6 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$