

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 $\mathbf{u}, \mathbf{v}$ in V and a scalar c , there are unique vectors $\mathbf{u} + \mathbf{v}$ and $c\mathbf{u}$ in V such that the following properties are satisfied.

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

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

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

For instance,

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

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 $\mathbf{u}, \mathbf{v} \in H$ and scalar $c \in \mathbb{R}$, we have

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

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

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 $A\mathbf{x} = \mathbf{0}$ is a subspace of $\mathbb{R}^n$. However, if $\mathbf{b} \neq \mathbf{0}$, the set of solutions of the system $A\mathbf{x} = \mathbf{b}$ is *not* a subspace of $\mathbb{R}^n$.
- (c) For any vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ in $\mathbb{R}^n$, the span $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ is a subspace of $\mathbb{R}^n$.
- (d) For any linear transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$, the image

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

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

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

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 $\text{Nul } A$, is the space of solutions of the linear system $A\mathbf{x} = \mathbf{0}$, that is,

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

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

$$\text{Col } A = \text{Span}\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}.$$

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

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

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

3 Linear transformations

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

- (a) $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$,
- (b) $T(c\mathbf{u}) = cT(\mathbf{u})$.

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

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

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 $F : \mathbf{M}(2, 2) \rightarrow \mathbf{M}(2, 2)$ 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 $D : \mathbf{P}_3 \rightarrow \mathbf{P}_2$, defined by

$$D(a_0 + a_1t + a_2t^2 + a_3t^3) = a_1 + 2a_2t + 3a_3t^2,$$

is a linear transformation.

Proposition 3.1. Let $T : V \rightarrow W$ 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 $T : V \rightarrow W$ be a linear transformation. Given vectors $v_1, v_2, \dots, v_k$ in V .

- (a) If $v_1, v_2, \dots, v_k$ are linearly dependent, then $T(v_1), T(v_2), \dots, T(v_k)$ are linearly dependent;
- (b) If $T(v_1), T(v_2), \dots, T(v_k)$ are linearly independent, then $v_1, v_2, \dots, v_k$ are linearly independent.

4 Independent sets and bases

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

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

we have

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

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

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

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

Theorem 4.2. Vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ ($k \geq 2$) 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

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

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

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k = \mathbf{0}.$$

Let $c_i \neq 0$. Then

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

□

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

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

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

then

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

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

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

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

Definition 4.4. Let H be a subspace of a vector space V . An ordered set $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$ 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, $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$.

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 $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$.

- (a) If some $\mathbf{v}_k$ is a linear combination of the other vectors of $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$, then

$$H = \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}, \mathbf{v}_{k+1}, \dots, \mathbf{v}_p\}.$$

- (b) If $H \neq \{0\}$, then some subsets of $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$ are bases for H .

Proof. It is clear that $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}, \mathbf{v}_{k+1}, \dots, \mathbf{v}_p\}$ is contained in H . Write

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

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

$$\begin{aligned} \mathbf{v} &= (a_1 + a_k c_1) \mathbf{v}_1 + \dots + (a_{k-1} + a_k c_{k-1}) \mathbf{v}_{k-1} \\ &\quad + (a_{k+1} + a_k c_{k+1}) \mathbf{v}_{k+1} + \dots + (a_p + a_k c_p) \mathbf{v}_p. \end{aligned}$$

This means that $\mathbf{v}$ is contained in $\text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k-1}, \mathbf{v}_{k+1}, \dots, \mathbf{v}_p\}$. □

Example 4.2. The vector space $\mathbf{P}(t)$ of polynomials of degree ≤ 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 2×2 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

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

Its reduced row echelon form is the matrix

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

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

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

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

$$\begin{aligned} \mathbf{b}_2 &= 4\mathbf{b}_1 & \iff & \mathbf{a}_2 = 4\mathbf{a}_1 \\ \mathbf{b}_4 &= 2\mathbf{b}_1 - \mathbf{b}_3 & \iff & \mathbf{a}_4 = 2\mathbf{a}_1 - \mathbf{a}_3 \\ \mathbf{b}_5 &= 4\mathbf{b}_1 + 2\mathbf{b}_3 & \iff & \mathbf{a}_5 = 4\mathbf{a}_1 + 2\mathbf{a}_3. \end{aligned}$$

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

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

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

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

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

6 Coordinate systems

Lecture 17

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

$$\mathbf{v} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n.$$

Proof. Trivial. □

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

$$\mathbf{v} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n.$$

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

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

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

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

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 $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p\}$ be a basis of a subspace H of $\mathbb{R}^n$. Let $P_{\mathcal{B}}$ be the matrix

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

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

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

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

Proof. Let $\mathbf{v} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_p \mathbf{v}_p$. Then

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

□

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

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

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

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

$$\mathbf{v} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n,$$

$$\mathbf{w} = d_1 \mathbf{v}_1 + d_2 \mathbf{v}_2 + \dots + d_n \mathbf{v}_n,$$

then

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

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

Thus

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

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

So the coordinate transformation is a linear transformation.

Now for any vector

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

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

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

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

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 $F : \mathbf{P}_3 \rightarrow \mathbb{R}^4$, defined by

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

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

Example 6.3. The vector space $\mathbf{M}(2, 2)$ of 2×2 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 $F : \mathbf{M}(2, 2) \rightarrow \mathbb{R}^4$, 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 $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a basis of a vector space V . Then any set of V consisting more than n vectors are linearly dependent.

Proof. Let $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p\}$ 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 $[\mathbf{u}_1]_{\mathcal{B}}, [\mathbf{u}_2]_{\mathcal{B}}, \dots, [\mathbf{u}_p]_{\mathcal{B}}$ of $\mathbb{R}^n$ must be linearly dependent. Then there exist constants $c_1, c_2, \dots, c_p$, not all zero, such that

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

Thus

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

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

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

This means that the vectors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_p$ are linearly dependent by definition. $\square$

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

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

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

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

- (a) If $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is linearly independent, then $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ is a basis of V .

(b) If $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} = V$, then $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ 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 $\text{Span } S' = V$. So S' is a basis of V ; therefore $\#(S') \geq n$, contradict to $\#(S') < n$. $\square$

8 Rank

Theorem 8.1. For any rectangular matrix A ,

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

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 $T : V \rightarrow W$, the **rank** of T is the dimension of the subspace $T(V)$.

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

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

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 $A\mathbf{x} = \mathbf{0}$. It is clear that

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

$\square$

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

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

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

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

9 Matrices of linear transformations

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

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

writing more compactly,

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

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

$$A = \left[[T(\mathbf{v}_1)]_{\mathcal{C}}, [T(\mathbf{v}_2)]_{\mathcal{C}}, \dots, [T(\mathbf{v}_n)]_{\mathcal{C}} \right].$$

Consider the isomorphisms

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

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

$$\mathbf{v} = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_n \mathbf{v}_n = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] \mathbf{x}.$$

Then

$$\begin{aligned} T(\mathbf{v}) &= T(x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \dots + x_n \mathbf{v}_n) \\ &= x_1 T(\mathbf{v}_1) + x_2 T(\mathbf{v}_2) + \dots + x_n T(\mathbf{v}_n) \\ &= [T(\mathbf{v}_1), T(\mathbf{v}_2), \dots, T(\mathbf{v}_n)] \mathbf{x} \\ &= [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m] A \mathbf{x}. \end{aligned}$$

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

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

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 $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $T(\mathbf{x}) = A\mathbf{x}$. Let $\mathcal{E}_n = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ be the standard basis of $\mathbb{R}^n$, and let $\mathcal{E}_m = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m\}$ be the standard basis of $\mathbb{R}^m$. Then

$$[T(\mathbf{e}_1), T(\mathbf{e}_2), \dots, T(\mathbf{e}_n)] = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m] A.$$

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

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

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

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

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

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

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

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

which is equivalent to

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

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

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

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

$$[\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m | T(\mathbf{v}_1), T(\mathbf{v}_2), \dots, T(\mathbf{v}_n)] \sim [I_m | A],$$

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 $T : \mathbb{R}^4 \rightarrow \mathbb{R}^3$ be defined by

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

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

(b) Verify that

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

is basis of V , and

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

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

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

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

$$\begin{aligned} [\mathbf{w}_1, \mathbf{w}_2 | T(\mathbf{v}_1), T(\mathbf{v}_2), T(\mathbf{v}_3)] &= \left[\begin{array}{cc|ccc} 0 & 1 & 2 & 3 & 3 \\ 1 & -1 & -1 & -1 & 0 \\ -2 & 1 & 0 & -1 & -3 \end{array} \right] \sim \\ &\left[\begin{array}{cc|ccc} 1 & -1 & -1 & -1 & 0 \\ 0 & 1 & 2 & 3 & 3 \\ -2 & 1 & 0 & -1 & -3 \end{array} \right] \sim \left[\begin{array}{cc|ccc} 1 & 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 3 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]. \end{aligned}$$

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 $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. A linear transformation $T : V \rightarrow V$ is called a **linear operator** on V . The matrix

$$A = \left[[T(\mathbf{v}_1)]_{\mathcal{B}}, [T(\mathbf{v}_2)]_{\mathcal{B}}, \dots, [T(\mathbf{v}_n)]_{\mathcal{B}} \right]$$

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

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

$$[\mathbf{v}'_1, \mathbf{v}'_2, \dots, \mathbf{v}'_n] = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]P.$$

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

$$[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n]P,$$

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

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

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

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

Find the matrix of T relative to the basis

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

Solution. Since

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

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

then

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

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

Let $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$. Then $[\mathbf{v}_1, \mathbf{v}_2] = [\mathbf{e}_1, \mathbf{e}_2]P$, where $P = [\mathbf{v}_1, \mathbf{v}_2] = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$. 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 $\begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ 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 $T : V \rightarrow V$, 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 $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ and $\mathcal{C} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be bases of a vector space V . Then

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

We may write this more compactly as

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

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

$$[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k] = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k] P.$$

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

$$\mathbf{v} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k] \mathbf{x} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k] \mathbf{y}.$$

Then

$$[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k] \mathbf{x} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k] P \mathbf{y}.$$

It follows that

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

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

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

Then

$$B \mathbf{p}_1 = \mathbf{v}_1, \quad B \mathbf{p}_2 = \mathbf{v}_2, \quad \dots, \quad B \mathbf{p}_k = \mathbf{v}_k.$$

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

$$B \mathbf{x} = \mathbf{v}_1, \quad B \mathbf{x} = \mathbf{v}_2, \quad \dots, \quad B \mathbf{x} = \mathbf{v}_k$$

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

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

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 $\mathbf{v}$ 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 $\mathbf{v}$ relative to the basis $\mathcal{C}$.

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

$$\begin{aligned} & \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 6 & 5 & 3 \\ 2 & 1 & 2 & 6 & 7 & 0 \\ 1 & 0 & 1 & 2 & 3 & -1 \\ 2 & 1 & 0 & 4 & 3 & 2 \\ 0 & 2 & 1 & 5 & 4 & 3 \end{array} \right] \begin{array}{l} R_2 - 2R_1 \\ R_3 - R_1 \\ \rightarrow \\ R_4 - 2R_1 \end{array} \\ & \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 6 & 5 & 3 \\ 0 & -3 & 0 & -6 & -3 & -6 \\ 0 & -2 & 0 & -4 & -2 & -4 \\ 0 & -3 & -2 & -8 & -7 & -4 \\ 0 & 2 & 1 & 5 & 4 & 3 \end{array} \right] \begin{array}{l} R_2/(-3) \\ \rightarrow \\ R_3 \leftrightarrow R_5 \end{array} \\ & \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 6 & 5 & 3 \\ 0 & 1 & 0 & 2 & 1 & 2 \\ 0 & 2 & 1 & 5 & 4 & 3 \\ 0 & -3 & -2 & -8 & -7 & -4 \\ 0 & -2 & 0 & -4 & -2 & -4 \end{array} \right] \begin{array}{l} R_3 - 2R_2 \\ R_4 + 3R_2 \\ \rightarrow \\ R_5 + 2R_2 \end{array} \\ & \left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 6 & 5 & 3 \\ 0 & 1 & 0 & 2 & 1 & 2 \\ 0 & 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & -2 & -2 & -4 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] \begin{array}{l} R_4 + 2R_3 \\ R_1 - R_3 \\ \rightarrow \\ R_1 - 2R_2 \end{array} \\ & \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 2 & 1 & 2 \\ 0 & 0 & 1 & 1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] \end{aligned}$$

Then $B = CP$, where

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

So the coordinate vector of $\mathbf{v}$ 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 $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be defined by

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

Consider two bases

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

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

Then

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

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

$$[\mathbf{v}'_1, \mathbf{v}'_2, \dots, \mathbf{v}'_n] = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]P.$$

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

$$[\mathbf{w}'_1, \mathbf{w}'_2, \dots, \mathbf{w}'_m] = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m]Q.$$

Let $T : V \rightarrow W$ 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

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

On the other hand, since

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

then

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

Thus

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

Therefore $QA' = AP$, that is,

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

□

Example 11.3. Let $T : \mathbf{P}_3 \rightarrow \mathbf{P}_2$ 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

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

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

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

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

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

One verifies that

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