CHAPTER 2

LINEAR TRANSFORMATIONS

2.1 Basic properties of linear transformation

Definition 2.1.1. Let V and W be vector spaces over F. A map $T : V \to W$ is said to be a *homomorphism* or *linear transformation* if for all $x, y \in V$ and $\alpha \in F$

(1) $T(x+y) = T(x) + T(y)$;
(2) $T(\alpha x) = \alpha T(x)$.

A vector space homomorphism T is also called a *linear map* or a *linear transformation* or a *linear operator* or simply *linear*.

It is easy to see that the two conditions for a linear transformation can be combined into a single condition as

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y), \quad \forall\, x, y \in V,\ \forall\, \alpha, \beta \in F$$

or

$$T(x + \alpha y) = T(x) + \alpha T(y), \quad \forall\, x, y \in V,\ \forall\, \alpha \in F.$$

Geometrically speaking, the linearity means the requirement for a straight line to be transformed into a straight line, since $x + \alpha y$ represents a straight line through x in the direction of y in V, and its image $T(x) + \alpha T(y)$ also represents a straight line through $T(x)$ in the direction of $T(y)$ in W.

Example 2.1.2 (Linear or not)**.** Consider the following functions:

(1) $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 2x$;
(2) $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = x^2 - x$;
(3) $h : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $h(x, y) = (x - y, 2x)$;
(4) $k : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $k(x, y) = (xy, x^2 + 1)$.

One can easily see that g and k are not linear, while f and h are linear. Further, on the 1-space $\mathbb{R}$, all polynomials of degree greater than one cannot be linear.

Example 2.1.3 (A matrix A as a linear transformation). (1) For an $m \times n$ matrix A, the transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by the matrix product

$$T(x) = Ax$$

is a linear transformation because by the distributive law $A(x + \alpha y) = Ax + \alpha Ay$ for any $x, y \in \mathbb{R}^n$ and for any scalar $\alpha \in \mathbb{R}$, and so $T(x + \alpha y) = T(x) + \alpha T(y)$. Therefore, a matrix A, identified with the transformation T, may be considered as a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$.

(2) For a vector space V, the *identity transformation* $id : V \to V$ is defined by $id(x) = x$ for all $x \in V$.
For vectors spaces V and W, the *zero transformation* $T_0 : V \to W$ is defined by $T_0(x) = \mathbf{0}$ (the zero vector) for all $x \in V$. Clearly, both transformations are linear.

The following theorem follows directly from the definition.

Theorem 2.1.4. *Let $T : V \to W$ be a linear transformation. Then*
(1) $T(\mathbf{0}) = \mathbf{0}$.
(2) For any $x_1, x_2, \ldots, x_n \in V$ and scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$,

$$T(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n) = \alpha_1 T(x_1) + \alpha_2 T(x_2) + \cdots + \alpha_n T(x_n).$$

Proof. **(1)** $T(0) = T(0 + 0) = T(0) + T(0) \Rightarrow T(0) = 0$.

OR

Since $\mathbf{0} = 0 \cdot \mathbf{0}$ in V, by scalar multiplication of T:

$$T(\mathbf{0}) = T(0 \cdot \mathbf{0}) = 0 \cdot T(\mathbf{0}) = \mathbf{0}.$$

(2) We proceed by induction on n. For $n = 1$, clearly $T(\alpha_1 x_1) = \alpha_1 T(x_1)$ holds by linearity. Assume the result holds for $n = m$. For $n = m + 1$:

$$\begin{aligned} T\left(\sum_{i=1}^{m+1} \alpha_i x_i\right) &= T\left(\sum_{i=1}^{m} \alpha_i x_i + \alpha_{m+1} x_{m+1}\right) \\ &= T\left(\sum_{i=1}^{m} \alpha_i x_i\right) + T(\alpha_{m+1} x_{m+1}) \quad \text{(by additivity)} \\ &= \sum_{i=1}^{m} \alpha_i T(x_i) + \alpha_{m+1} T(x_{m+1}) \quad \text{(by induction hypothesis and linearity)} \\ &= \sum_{i=1}^{m+1} \alpha_i T(x_i). \end{aligned}$$

Hence, the statement holds for all $n \geq 1$. □

Nontrivial and important examples of linear transformations in a geometry are the rotations, reflections, and projections.

Example 2.1.5 (The rotations, reflections and projections in a geometry)**.** (1) Let θ denote the angle between the x-axis and a fixed vector in $\mathbb{R}^2$. Then the matrix

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

defines a linear transformation on $\mathbb{R}^2$ that rotates any vector in $\mathbb{R}^2$ through the angle θ about the origin. It is called a *rotation* by the angle θ.

(2) The *projection* on the x-axis is the linear transformation $P : \mathbb{R}^2 \to \mathbb{R}^2$ defined by, for $\mathbf{x} = (x, y) \in \mathbb{R}^2$,

$$P(\mathbf{x}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}.$$

(3) The linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by, for $\mathbf{x} = (x, y)$,

$$T(\mathbf{x}) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}$$

is the *reflection* about the x-axis.

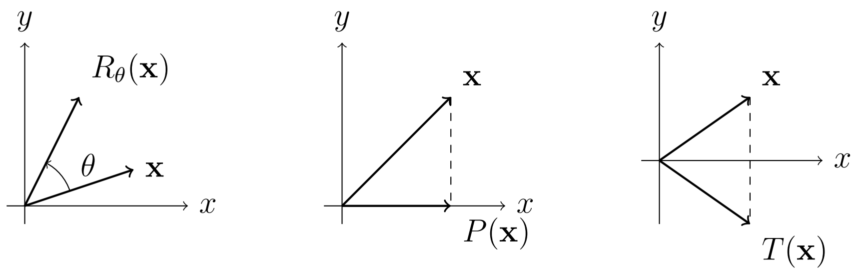

Figure 2.1: Three linear transformations on $\mathbb{R}^2$

Exercise 2.1.6. Find the matrix of the reflection about the line $y = x$ in the plane $\mathbb{R}^2$.

Example 2.1.7 (Differentiations and integrations in calculus)**.** In calculus, it is well known that two transformations

$$D : P_n(\mathbb{R}) \to P_{n-1}(\mathbb{R}), \quad \mathcal{I} : P_{n-1}(\mathbb{R}) \to P_n(\mathbb{R})$$

defined by differentiation and integration,

$$D(f)(x) = f'(x), \quad \mathcal{I}(f)(x) = \int_0^x f(t)dt,$$

satisfy linearity, and so they are linear transformations.

Definition 2.1.8. Let V and W be two vector spaces, and let $T : V \to W$ be a linear transformation from V into W.

(1) $\mathrm{Ker}(T) = \{v \in V : T(v) = 0\} \subseteq V$ is called the *kernel* of T.

(2) $\mathrm{Im}(T) = \{T(v) \in W : v \in V\} = T(V) \subseteq W$ is called the *image* of T.

Example 2.1.9. Let V and W be vector spaces and let $id : V \to V$ and $T_0 : V \to W$ be the identity and the zero transformations, respectively. Then it is easy to see that $\mathrm{Ker}(id) = \{0\}$, $\mathrm{Im}(id) = V$, $\mathrm{Ker}(T_0) = V$ and $\mathrm{Im}(T_0) = \{0\}$.

Theorem 2.1.10. *Let $T : V \to W$ be a linear transformation from a vector space V to a vector space W. Then the kernel* $\mathrm{Ker}(T)$ *and the image* $\mathrm{Im}(T)$ *are subspaces of V and W, respectively.*

Proof. **Nonempty:** Since $T(0) = 0$, each of $\mathrm{Ker}(T)$ and $\mathrm{Im}(T)$ is nonempty as both of them contain the 0 vector.

(1) Kernel is a subspace: For any $x, y \in \mathrm{Ker}(T)$ and for any scalar α,

$$T(x + \alpha y) = T(x) + \alpha T(y) = 0 + \alpha \cdot 0 = 0.$$

Hence $x + \alpha y \in \mathrm{Ker}(T)$, so that $\mathrm{Ker}(T)$ is a subspace of V.

(2) Image is a subspace: If $v, w \in \mathrm{Im}(T)$, then by the definition of image set, there exist x and y in V such that $T(x) = v$ and $T(y) = w$. Thus, for any scalar α,

$$v + \alpha w = T(x) + \alpha T(y) = T(x + \alpha y).$$

Since x and y are in V and V is a vector space, $x + \alpha y$ is also in V. Therefore, $T(x + \alpha y) \in \mathrm{Im}(T)$, the image of T. Thus $v + \alpha w \in \mathrm{Im}(T)$, so that $\mathrm{Im}(T)$ is a subspace of W. □

Example 2.1.11 ($\mathrm{Ker}(A) = \mathcal{N}(A)$ and $\mathrm{Im}(A) = \mathcal{C}(A)$ for any matrix A)**.** Let $A : \mathbb{R}^n \to \mathbb{R}^m$ be the linear transformation defined by an $m \times n$ matrix A as in Example 2.1.3 (1).

The kernel $\mathrm{Ker}(A)$ of A consists of all solutions of the homogeneous system $Ax = 0$. Hence, the kernel $\mathrm{Ker}(A)$ of A is nothing but the null space $\mathcal{N}(A)$ of the matrix A, and the image $\mathrm{Im}(A)$ of A is just the column space $\mathcal{C}(A) = \mathrm{Im}(A) \subseteq \mathbb{R}^m$ of the matrix A. Recall that Ax is a linear combination of the column vectors of A.

Example 2.1.12 (The trace is linear)**.** A *trace* is a function $\mathrm{tr} : M_{n\times n}(\mathbb{R}) \to \mathbb{R}$ defined by the sum of diagonal entries:

$$\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii}$$

for $A = [a_{ij}] \in M_{n\times n}(\mathbb{R})$, and $\mathrm{tr}(A)$ is called the *trace* of the matrix A. It is easy to see that

$$\mathrm{tr}(A + B) = \mathrm{tr}(A) + \mathrm{tr}(B) \quad \text{and} \quad \mathrm{tr}(\alpha A) = \alpha\,\mathrm{tr}(A)$$

for any two matrices A and B in $M_{n\times n}(\mathbb{R})$, which means that 'tr' is a linear transformation from $M_{n\times n}(\mathbb{R})$ to $\mathbb{R}$.

Also, one can easily show that the set of all $n \times n$ matrices with trace 0 is a subspace of the vector space $M_{n\times n}(\mathbb{R})$. □

Exercise 2.1.13. Let $W = \{A \in M_{n\times n}(\mathbb{R}) : \mathrm{tr}(A) = 0\}$. Show that W is a subspace, and find a basis for W.

Exercise 2.1.14. Show that, for any matrices A and B in $M_{n\times n}(\mathbb{R})$, $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

One of the most significant properties of linear transformations is that they are completely determined by their values on a basis.

Theorem 2.1.15. *Let V and W be vector spaces. Let $\{v_1, v_2, \dots, v_n\}$ be a basis for V and let $w_1, w_2, \dots, w_n$ be any vectors (possibly repeated) in W. Then there exists a unique linear transformation $T : V \to W$ such that $T(v_i) = w_i$ for $i = 1, 2, \dots, n$.*

Proof. Let $x \in V$. Then it has a unique expression $x = \sum_{i=1}^n a_i v_i$ for some scalars $a_1, a_2, \dots, a_n$. Define

$$T : V \to W \quad \text{by} \quad T(x) = \sum_{i=1}^n a_i w_i.$$

In particular, $T(v_i) = w_i$ for $i = 1, 2, \dots, n$.

Linearity: For $x = \sum_{i=1}^n a_i v_i$, $y = \sum_{i=1}^n b_i v_i \in V$ and α a scalar, we have $x + \alpha y = \sum_{i=1}^n (a_i + \alpha b_i) v_i$. Then

$$\begin{aligned} T(x + \alpha y) &= \sum_{i=1}^n (a_i + \alpha b_i) w_i \\ &= \sum_{i=1}^n a_i w_i + \alpha \sum_{i=1}^n b_i w_i \\ &= T(x) + \alpha T(y). \end{aligned}$$

Uniqueness: Suppose that $S : V \to W$ is linear and $S(v_i) = w_i$ for $i = 1, 2, \dots, n$. Then for any $x \in V$ with $x = \sum_{i=1}^n a_i v_i$, we have

$$S(x) = \sum_{i=1}^n a_i S(v_i) = \sum_{i=1}^n a_i w_i = T(x).$$

Hence, we have $S = T$. □

The uniqueness in the above theorem can be stated as follows.

Corollary 2.1.16. *Let V and W be vector spaces and let $\{v_1, v_2, \dots, v_n\}$ be a basis for V. If $S, T : V \to W$ are linear transformations and $S(v_i) = T(v_i)$ for $i = 1, 2, \dots, n$; then $S = T$, i.e., $S(x) = T(x)$ for all $x \in V$.*

Example 2.1.17 (Linear extension of a transformation defined on a basis)**.** Let $w_1 = (1, 0)$, $w_2 = (2, -1)$, $w_3 = (4, 3)$ be three vectors in $\mathbb{R}^2$.

(1) Let $B = \{e_1, e_2, e_3\}$ be the standard basis for the 3-space $\mathbb{R}^3$, and let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation defined by

$$T(e_1) = w_1, \quad T(e_2) = w_2, \quad T(e_3) = w_3.$$

Find a formula for $T(x_1, x_2, x_3)$, and then use it to compute $T(2, -3, 5)$.

(2) Let $S = \{v_1, v_2, v_3\}$ be another basis for $\mathbb{R}^3$, where $v_1 = (1, 1, 1)$, $v_2 = (1, 1, 0)$, $v_3 = (1, 0, 0)$, and let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear transformation defined by

$$T(v_1) = w_1, \quad T(v_2) = w_2, \quad T(v_3) = w_3.$$

Find a formula for $T(x_1, x_2, x_3)$, and then use it to compute $T(2, -3, 5)$.

Solution. (1) For $x = (x_1, x_2, x_3) = x_1e_1 + x_2e_2 + x_3e_3 \in \mathbb{R}^3$,

$$\begin{aligned} T(x) &= \sum_{i=1}^{3} x_i T(e_i) = \sum_{i=1}^{3} x_i w_i \\ &= x_1(1,0) + x_2(2,-1) + x_3(4,3) \\ &= (x_1 + 2x_2 + 4x_3,\ -x_2 + 3x_3). \end{aligned}$$

Thus, $T(2,-3,5) = (16, 18)$. In matrix notation, this can be written as

$$\begin{bmatrix} 1 & 2 & 4 \\ 0 & -1 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 + 4x_3 \\ -x_2 + 3x_3 \end{bmatrix}.$$

(2) In this case, we need to express $x = (x_1, x_2, x_3)$ as a linear combination of v_1, v_2, v_3, i.e.,

$$\begin{aligned} (x_1, x_2, x_3) &= \sum_{i=1}^{3} \alpha_i v_i = \alpha_1(1,1,1) + \alpha_2(1,1,0) + \alpha_3(1,0,0) \\ &= (\alpha_1 + \alpha_2 + \alpha_3)e_1 + (\alpha_1 + \alpha_2)e_2 + \alpha_1 e_3. \end{aligned}$$

By equating corresponding components we obtain a system of equations:

$$\begin{cases} \alpha_1 + \alpha_2 + \alpha_3 &= x_1 \\ \alpha_1 + \alpha_2 &= x_2 \\ \alpha_1 &= x_3. \end{cases}$$

The solution is $\alpha_1 = x_3$, $\alpha_2 = x_2 - x_3$, $\alpha_3 = x_1 - x_2$. Therefore,

$$\begin{aligned} (x_1, x_2, x_3) &= x_3 v_1 + (x_2 - x_3)v_2 + (x_1 - x_2)v_3, \quad \text{and} \\ T(x_1, x_2, x_3) &= x_3 T(v_1) + (x_2 - x_3)T(v_2) + (x_1 - x_2)T(v_3) \\ &= x_3(1,0) + (x_2 - x_3)(2,-1) + (x_1 - x_2)(4,3) \\ &= (4x_1 - 2x_2 - x_3,\ 3x_1 - 4x_2 + x_3). \end{aligned}$$

From this formula, we obtain $T(2,-3,5) = (9, 23)$.

□

Exercise 2.1.18. Is there a linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$ such that $T(3,1,0) = (1,1)$ and $T(-6,-2,0) = (2,1)$? If yes, can you find an expression of $T(x)$ for $x = (x_1, x_2, x_3)$ in $\mathbb{R}^3$?

Exercise 2.1.19. Let V and W be vector spaces and $T : V \to W$ be linear. Let $\{w_1, w_2, \ldots, w_k\}$ be a linearly independent subset of the image $\mathrm{Im}(T) \subseteq W$. Suppose that $\alpha = \{v_1, v_2, \ldots, v_k\}$ is chosen so that $T(v_i) = w_i$ for $i = 1, 2, \ldots, k$. Prove that α is linearly independent.

2.2 Invertible linear transformations

A function f from a set X to a set Y is said to be *invertible* if there is a function g from Y to X such that their compositions satisfy $g \circ f = id$ and $f \circ g = id$. Such a function g is called the *inverse* function of f and denoted by $g = f^{-1}$. One can notice that if there exists an invertible function from a set X into another set Y, then it gives a one-to-one correspondence between these two sets so that they can be identified as sets. A useful criterion for a function between two given sets to be invertible is that it is one-to-one and onto.

Lemma 2.2.1. *A function $f : X \to Y$ is invertible if and only if it is bijective (i.e., one-to-one and onto).*

Proof. Suppose $f : X \to Y$ is invertible, and let $g : Y \to X$ be its inverse.

$$\begin{aligned} & f(u) = f(v) \\ \Rightarrow\ & g(f(u)) = g(f(v)) \\ \Rightarrow\ & u = v. \end{aligned}$$

Thus, f is one-to-one.

For each $y \in Y$, let $g(y) = x$ in X. Then $f(x) = f(g(y)) = y$. Thus, f is onto.

Conversely, suppose f is bijective. Then, for each $y \in Y$, there is a unique $x \in X$ such that $f(x) = y$. Now, for each $y \in Y$ define $g(y) = x$. Then one can easily check that $g : Y \to X$ is well defined. Further, for each $x \in X$,

$$(g \circ f)(x) = g(f(x)) = g(y) = x = id(x).$$

Thus, $f \circ g = id$ and similarly, we have $g \circ f = id$. Hence, g is the inverse function of f. □

If $T : V \to W$ and $S : W \to Z$ are linear transformations, then it is easy to show that their composition $(S \circ T)(v) = S(T(v))$ is also a linear transformation. In particular, if two linear transformations are defined by matrices $A : \mathbb{R}^n \to \mathbb{R}^m$ and $B : \mathbb{R}^m \to \mathbb{R}^k$ as in Example 2.1.3 (1), then their composition is nothing but the matrix product BA of them, i.e., $(B \circ A)(x) = B(Ax) = (BA)x$.

The following lemma shows that if a given function is an invertible linear transformation from a vector space into another, then the linearity is preserved by the inversion. That is, inverse of a linear transformation is also a linear transformation.

Lemma 2.2.2. *Let V and W be vector spaces over F. If $T : V \to W$ is an invertible linear transformation, then its inverse $T^{-1} : W \to V$ is also linear.*

Proof. Let $w_1, w_2 \in W$, and let $\alpha \in F$ be any scalar. Since T is invertible (i.e., bijective), there exist unique vectors v_1 and v_2 in V such that $T(v_1) = w_1$ and $T(v_2) = w_2$. Then

$$\begin{aligned} T^{-1}(w_1 + \alpha w_2) &= T^{-1}(T(v_1) + \alpha T(v_2)) \\ &= T^{-1}(T(v_1 + \alpha v_2)) && \text{(by linearity of } T\text{)} \\ &= v_1 + \alpha v_2 \\ &= T^{-1}(w_1) + \alpha T^{-1}(w_2). \end{aligned}$$

□

Definition 2.2.3. A linear transformation $T : V \to W$ from a vector space V to another W is called an *isomorphism* if it is invertible (i.e., one-to-one and onto). In this case, we say that V and W are *isomorphic* to each other.

Example 2.2.4 (The vector space $P_2(\mathbb{R})$ is isomorphic to $\mathbb{R}^{2+1}$)**.** Consider the vector space $P_2(\mathbb{R}) = \{a + bx + cx^2 : a, b, c \in \mathbb{R}\}$ of all polynomials of degree ≤ 2 with real coefficients. To each polynomial $a + bx + cx^2$ in the space $P_2(\mathbb{R})$, one can assign the column vector $[a\ b\ c]^T$ in $\mathbb{R}^3$, i.e., $(a, b, c) \in \mathbb{R}^3$. Then it can be easily seen that (**verify!**) it is an isomorphism from the vector space $P_2(\mathbb{R})$ to the 3-space $\mathbb{R}^3$, by which one can *identify* the polynomial $a + bx + cx^2$ with the column vector $[a\ b\ c]^T$.

In general, the vector space $P_n(\mathbb{R})$ can be identified with the $(n + 1)$-space $\mathbb{R}^{n+1}$.

It is clear from Lemma 2.2.2 that if T is an isomorphism, then its inverse T^{-1} is also an isomorphism with $(T^{-1})^{-1} = T$. In particular, if a linear transformation $A : \mathbb{R}^n \to \mathbb{R}^n$ is defined by an invertible $n \times n$ matrix A as in Example 2.1.3 (1), then the inverse matrix A^{-1} plays the inverse linear transformation, so that it is also an isomorphism on $\mathbb{R}^n$. That is, a linear transformation $A : \mathbb{R}^n \to \mathbb{R}^n$ defined by an $n \times n$ square matrix A is an isomorphism if and only if A is invertible, that is, $\operatorname{rank} A = n$.

Exercise 2.2.5. Suppose that S and T are linear transformations whose composition $S \circ T$ is well defined. Prove that

(1) if $S \circ T$ is one-to-one, so is T.
(2) if $S \circ T$ is onto, so is S.
(3) if S and T are isomorphisms, so is $S \circ T$.
(4) if A and B are two $n \times n$ matrices of rank n, so is AB.

Exercise 2.2.6. Let $T : V \to W$ be a linear transformation. Prove that

(1) T is one-to-one if and only if $\operatorname{Ker}(T) = \{0\}$.
(2) If $V = W$, then T is one-to-one if and only if T is onto.

Theorem 2.2.7. *Two vector spaces V and W over F are isomorphic if and only if* $\dim V = \dim W$.

Proof. Let $T : V \to W$ be an isomorphism, and let $\{v_1, v_2, \ldots, v_n\}$ be a basis for V. Then we show that the set $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ is a basis for W so that $\dim W = n = \dim V$.

Claim. $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ **is linearly independent.**

$$
\begin{aligned}
&\alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n) = 0 && \\
\Rightarrow\ & T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = 0 && (\because T \text{ is linear}) \\
\Rightarrow\ & \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0 && (\because T \text{ is one-one}) \\
\Rightarrow\ & \alpha_i = 0,\ \forall\, i = 1, 2, \ldots, n && (\because v_1, \ldots, v_n \text{ are LI}).
\end{aligned}
$$

Claim. $\{T(v_1), T(v_2), \ldots, T(v_n)\}$ **spans** W**.**
Since T is onto, for any $y \in W$ there exists an $x \in V$ such that $T(x) = y$. Since $\{v_1, v_2, \ldots, v_n\}$ is a basis for V, $x = \sum_{i=1}^n \alpha_i v_i$ for some $\alpha_i \in F$. Then

$$y = T(x) = T(\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n) = \alpha_1 T(v_1) + \alpha_2 T(v_2) + \cdots + \alpha_n T(v_n),$$

i.e., y is a linear combination of $T(v_1), T(v_2), \ldots, T(v_n)$.

Conversely, assume that $\dim V = \dim W = n$, and let $\{v_1, v_2, \ldots, v_n\}$ and $\{w_1, w_2, \ldots, w_n\}$ be bases for V and W, respectively. By Theorem 2.1.15, there exists unique linear transformations $T : V \to W$ such that $T(v_i) = w_i$ for $i = 1, 2, \ldots, n$.

Claim. T is one-one.
Let $T(x) = T(y)$ for some $x = \alpha_1 v_1 + \cdots + \alpha_n v_n$ and $y = \beta_1 v_1 + \cdots + \beta_n v_n$ in V. Then

$$\begin{aligned} & T(x) = T(y) \\ \Rightarrow\ & T(\alpha_1 v_1 + \cdots + \alpha_n v_n) = T(\beta_1 v_1 + \cdots + \beta_n v_n) \\ \Rightarrow\ & \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n) = \beta_1 T(v_1) + \cdots + \beta_n T(v_n) \\ \Rightarrow\ & \alpha_1 w_1 + \cdots + \alpha_n w_n = \beta_1 w_1 + \cdots + \beta_n w_n \\ \Rightarrow\ & (\alpha_1 - \beta_1) w_1 + \cdots + (\alpha_n - \beta_n) w_n = 0 \\ \Rightarrow\ & \alpha_i = \beta_i,\ \forall\, i && (\because w_1, \ldots, w_n \text{ are LI}) \\ \Rightarrow\ & x = y. \end{aligned}$$

Thus, T is one-one.

Claim. T is onto.
Let $w \in W$. Then can be written as a linear combination of basis for W as $w = \alpha_1 w_1 + \cdots + \alpha_n w_n$. Then the element $v = \alpha_1 v_1 + \cdots + \alpha_n v_n \in V$ as it is a linear combination of basis elements of V and

$$T(v) = T(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 T(v_1) + \cdots + \alpha_n T(v_n) = \alpha_1 w_1 + \cdots + \alpha_n w_n = w.$$

Thus, T is onto.

Hence, V and W are isomorphic. □

From the first part of the proof of the above theorem, we have the following corollary.

Corollary 2.2.8. *If T is an isomorphism of V onto W, prove that T maps a basis of V onto a basis of W.*

In particular, taking the vector space W over F to be $F^{(n)}$, we have the following corollary.

Corollary 2.2.9. *If V is a finite dimensional vector space over F then V is isomorphic to $F^{(n)}$ for some n.*

Corollary 2.2.10. *$F^{(n)}$ and $F^{(m)}$ are isomorphic if and only if $n = m$.*

Exercise 2.2.11. Let $\mathcal{V} = \{V : V \text{ is a vector space over } F\}$. Define a relation '$\sim$' on $\mathcal{V}$ as follows: for $U, V \in \mathcal{V}$, $U \sim V$ if U is isomorphic to V. Show that '$\sim$' is an equivalence relation, i.e., isomorphism of vector spaces is an equivalence relation.

Corollary 2.2.12. *Let V and W be vector spaces.*

(1) If $\dim V = n$, then V is isomorphic to the n-space $\mathbb{R}^n$.
(2) If $\dim V = \dim W$, any bijective function from a basis for V to a basis for W can be extended to an isomorphism from V to W.

Proof. Exercise. □

An isomorphism between a vector space V and $\mathbb{R}^n$ in the above corollary depends on the choices of bases for two spaces as shown in Theorem 2.2.7. However, an isomorphism is uniquely determined if we fix the bases in which the order of the vectors is also fixed.

An *ordered basis* for a vector space is a basis endowed with a specific order. For example, in $\mathbb{R}^3$, two bases $\{e_1, e_2, e_3\}$ with the order e_1, e_2, e_3 and $\{e_2, e_1, e_3\}$ with the order e_2, e_1, e_3 are clearly different as ordered bases, but the same as unordered ones. The basis $\{e_1, e_2, \dots, e_n\}$ with the order $e_1, e_2, \dots, e_n$ is called the *standard ordered basis* for $\mathbb{R}^n$. However, we often say simply a basis for an ordered basis if there is no ambiguity in the context.

Author Note: In what follows in this unit, we use the notations of our reference text although they may feel a bit unnatural to us.

Let V be a vector space of dimension n with an ordered basis $\alpha = \{v_1, v_2, \dots, v_n\}$, and let $\beta = \{e_1, e_2, \dots, e_n\}$ be the standard ordered basis for $\mathbb{R}^n$. Then the isomorphism $\Phi : V \to \mathbb{R}^n$ defined by $\Phi(v_i) = e_i$ is called the *natural isomorphism* with respect to the basis α. By this isomorphism, a vector in V can be identified with a column vector in $\mathbb{R}^n$. In fact, for any $x = \sum_{i=1}^n a_i v_i \in V$, the image of x under this natural isomorphism is written as

$$\Phi(x) = \sum_{i=1}^n a_i \Phi(v_i) = \sum_{i=1}^n a_i e_i = (a_1, \dots, a_n) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n,$$

which is called the *coordinate vector* of x with respect to the ordered basis α, and it is denoted by $[x]_\alpha$. Clearly $[v_i]_\alpha = e_i$.

Example 2.2.13. (1) Recall that, from Example 2.1.5, the rotation by the angle θ of $\mathbb{R}^2$ is given by the matrix

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Clearly, it is invertible and hence is an isomorphism of $\mathbb{R}^2$. In fact, the inverse R_θ^{-1} is another rotation $R_{-\theta}$.

(2) Let $\alpha = \{e_1, e_2\}$ be the standard ordered basis, and let $\beta = \{v_1, v_2\}$, where $v_i = R_\theta e_i$, $i = 1, 2$. Then β is also a basis for $\mathbb{R}^2$. The coordinate vectors of v_i with respect to α are

$$[v_1]_\alpha = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \quad [v_2]_\alpha = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix},$$

while

$$[v_1]_\beta = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad [v_2]_\beta = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

If we choose $\alpha' = \{e_2, e_1\}$ as a different ordered basis for $\mathbb{R}^2$, then the coordinate vectors of v_i with respect to α' are

$$[v_1]_{\alpha'} = \begin{bmatrix} \sin\theta \\ \cos\theta \end{bmatrix}, \quad [v_2]_{\alpha'} = \begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix}.$$

Exercise 2.2.14. Find the coordinate vector of $5 + 2x + 3x^2$ with respect to the given ordered basis α for $P_2(\mathbb{R})$:

(1) $\alpha = \{1,\ x,\ x^2\}$;
(2) $\alpha = \{1 + x,\ 1 + x^2,\ x + x^2\}$.

2.3 Matrices of linear transformations

Let $T : V \to W$ be a linear transformation from an n-dimensional vector space V to an m-dimensional vector space W, and let $\alpha = \{v_1, \ldots, v_n\}$ and $\beta = \{w_1, \ldots, w_m\}$ be any ordered bases for V and W, respectively, which will be fixed throughout this section. Then by Theorem 2.1.15, the linear transformation T is completely determined by its values on a basis α: Write them as

$$\begin{cases} T(v_1) = a_{11}w_1 + a_{21}w_2 + \cdots + a_{m1}w_m \\ T(v_2) = a_{12}w_1 + a_{22}w_2 + \cdots + a_{m2}w_m \\ \quad\vdots \\ T(v_n) = a_{1n}w_1 + a_{2n}w_2 + \cdots + a_{mn}w_m, \end{cases}$$

or, in a short form,

$$T(v_j) = \sum_{i=1}^{m} a_{ij}w_i \quad \text{for } 1 \le j \le n,$$

for some scalars a_{ij} ($i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n$).

Now, for any vector $x = \sum_{j=1}^{n} x_j v_j \in V$,

$$T(x) = \sum_{j=1}^{n} x_j T(v_j) = \sum_{j=1}^{n} x_j \sum_{i=1}^{m} a_{ij}w_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}x_j \right) w_i.$$

Equivalently, the coordinate vector of $T(x)$ with respect to the basis β in W is

$$[T(x)]_\beta = \begin{bmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}x_j \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = A[x]_\alpha.$$

That is, for any $x \in V$ the coordinate vector $[T(x)]_\beta$ of $T(x)$ in W is just the product of a fixed matrix A and the coordinate vector $[x]_\alpha$ of x. Note that

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} [T(v_1)]_\beta & \cdots & [T(v_n)]_\beta \end{bmatrix}$$

is the matrix whose column vectors are just the coordinate vectors $[T(v_j)]_\beta$ of $T(v_j)$ with respect to the basis β.

Definition 2.3.1. The matrix A is called the **associated matrix** for T (or the **matrix representation** of T) with respect to the ordered bases α and β, and denoted by $A = [T]_\alpha^\beta$. When $V = W$ and $\alpha = \beta$, we simply write $[T]_\alpha$ for $[T]_\alpha^\alpha$.

Now, the argument so far can be summarized in the following theorem.

Theorem 2.3.2. *Let $T : V \to W$ be a linear transformation from an n-dimensional vector space V to an m-dimensional vector space W. For fixed ordered bases $\alpha = \{v_1, v_2, \dots, v_n\}$ for V and β for W, there corresponds a unique associated $m \times n$ matrix $[T]_\alpha^\beta$ for T such that for any vector $x \in V$ the coordinate vector $[T(x)]_\beta$ of $T(x)$ with respect to β is given as a matrix product of the associated matrix $[T]_\alpha^\beta$ for T and the coordinate vector $[x]_\alpha$, i.e.,*

$$[T(x)]_\beta = [T]_\alpha^\beta [x]_\alpha.$$

The associated matrix $[T]_\alpha^\beta$ is given as

$$[T]_\alpha^\beta = \begin{bmatrix} [T(v_1)]_\beta & [T(v_2)]_\beta & \dots & [T(v_n)]_\beta \end{bmatrix}.$$

Let us consider the following examples for the computation of the associated matrices for linear transformations.

Example 2.3.3 (The associated matrix $[id]_\alpha$)**.** Let $id : V \to V$ be the identity transformation on a vector space V. Then for any ordered basis α for V, the matrix $[id]_\alpha = I$, the identity matrix, because if $\alpha = \{v_1, v_2, \dots, v_n\}$, then

$$\begin{cases} id(v_1) = 1v_1 + 0v_2 + \cdots + 0v_n \\ id(v_2) = 0v_1 + 1v_2 + \cdots + 0v_n \\ \quad\vdots \\ id(v_n) = 0v_1 + 0v_2 + \cdots + 1v_n. \end{cases}$$

Example 2.3.4 (The associated matrix $[T]_\alpha^\beta$)**.** Let $T : P_1(\mathbb{R}) \to P_2(\mathbb{R})$ be the linear transformation defined by

$$(T(p))(x) = xp(x).$$

Find the associated matrix $[T]_\alpha^\beta$ with respected to ordered bases $\alpha = \{1, x\}$ and $\beta = \{1, x, x^2\}$ for $P_1(\mathbb{R})$ and $P_2(\mathbb{R})$, respectively.

Solution: Clearly,

$$\begin{cases} (T(1))(x) = x = 0 \cdot 1 + 1x + 0x^2 \\ (T(x))(x) = x^2 = 0 \cdot 1 + 0x + 1x^2. \end{cases}$$

Hence, the associated matrix for T is the transpose of the coefficient matrix in this expression, that is,

$$[T]_\alpha^\beta = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Example 2.3.5 (The associated matrices $[T]_\alpha^\beta$ and $[T]_\alpha^{\beta'}$)**.** Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation defined by $T(x, y) = (x + 2y,\ 0,\ 2x + 3y)$ with respect to the standard bases

α and β for $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively. Then

$$\begin{cases} T(e_1) = T(1,0) = (1,0,2) = 1e_1 + 0e_2 + 2e_3 \\ T(e_2) = T(0,1) = (2,0,3) = 2e_1 + 0e_2 + 3e_3. \end{cases}$$

Hence,

$$[T]_\alpha^\beta = \begin{bmatrix} 1 & 2 \\ 0 & 0 \\ 2 & 3 \end{bmatrix}.$$

If $\beta' = \{e_3, e_2, e_1\}$, then

$$[T]_\alpha^{\beta'} = \begin{bmatrix} 2 & 3 \\ 0 & 0 \\ 1 & 2 \end{bmatrix}.$$

Example 2.3.6 (The associated matrix $[T]_\alpha$ for the standard basis α)**.** Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation given by $T(1,1) = (0,1)$ and $T(-1,1) = (2,3)$. Find the matrix representation $[T]_\alpha$ of T with respect to the standard basis $\alpha = \{e_1, e_2\}$.

Solution: Note $(a,b) = ae_1 + be_2$ for any $(a,b) \in \mathbb{R}^2$. Thus, the definition of T shows

$$\begin{aligned} T(e_1) + T(e_2) &= T(e_1 + e_2) = T(1,1) = (0,1) = e_2, \\ -T(e_1) + T(e_2) &= T(-e_1 + e_2) = T(-1,1) = (2,3) = 2e_1 + 3e_2. \end{aligned}$$

By solving these equations, we obtain

$$\begin{aligned} T(e_1) &= -e_1 - e_2, \\ T(e_2) &= e_1 + 2e_2. \end{aligned}$$

Therefore,

$$[T]_\alpha = \begin{bmatrix} -1 & 1 \\ -1 & 2 \end{bmatrix}.$$

Example 2.3.7 (The associated matrix $[T]_\beta$ for a non-standard basis β)**.** Let T be the linear transformation given in Example 4.15. Find $[T]_\beta$ for a basis $\beta = \{v_1, v_2\}$, where $v_1 = (0,1)$ and $v_2 = (2,3)$.

Solution: From Example 4.15,

$$\begin{aligned} T(v_1) &= \begin{bmatrix} -1 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} = [T(v_1)]_\alpha, \\ T(v_2) &= \begin{bmatrix} -1 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix} = [T(v_2)]_\alpha. \end{aligned}$$

To write these vectors as linear combinations of basis vectors in β, we put

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} = av_1 + bv_2 = \begin{bmatrix} 2b \\ a + 3b \end{bmatrix}, \quad \begin{bmatrix} 1 \\ 4 \end{bmatrix} = cv_1 + dv_2 = \begin{bmatrix} 2d \\ c + 3d \end{bmatrix}.$$

Solving for a, b, c and d, we obtain

$$[T(v_1)]_\beta = \begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad [T(v_2)]_\beta = \begin{bmatrix} c \\ d \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 5 \\ 1 \end{bmatrix}.$$

Therefore,

$$[T]_\beta = \frac{1}{2} \begin{bmatrix} 1 & 5 \\ 1 & 1 \end{bmatrix}.$$

Remark 2.3.8. Let V and W be vector spaces with bases α and β, respectively, and let $T : V \to W$ be a linear transformation with the matrix representation $[T]_\alpha^\beta = A$. Then it is clear that $\mathrm{Ker}(T)$ and $\mathrm{Im}(T)$ are isomorphic to the null space $\mathcal{N}(A)$ and the column space $\mathcal{C}(A)$, respectively, via the natural isomorphisms. In particular, if $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$ with the standard bases, then $\mathrm{Ker}(T) = \mathcal{N}(A)$ and $\mathrm{Im}(T) = \mathcal{C}(A)$. Therefore, from Theorem 1.6.5 (Rank Theorem), we have

$$\dim \mathrm{Ker}(T) + \dim \mathrm{Im}(T) = \dim V.$$

Exercise 2.3.9. Find the matrix representations $[T]_\alpha$ and $[T]_\beta$ of each of the following linear transformations T on $\mathbb{R}^3$ with respect to the standard basis $\alpha = \{e_1, e_2, e_3\}$ and another $\beta = \{e_3, e_2, e_1\}$:

(1) $T(x, y, z) = (2x - 3y + 4z,\ 5x - y + 2z,\ 4x + 7y)$,
(2) $T(x, y, z) = (2y + z,\ x - 4y,\ 3x)$.

Also, find the matrix representation $[T]_\alpha^\beta$ of each of the linear transformations T.

Exercise 2.3.10. Let $T : \mathbb{R}^4 \to \mathbb{R}^3$ be the linear transformation defined by

$$T(x, y, z, u) = (x + 2y,\ x - 3z + u,\ 2y + 3z + 4u).$$

Let α and β be the standard bases for $\mathbb{R}^4$ and $\mathbb{R}^3$, respectively. Find $[T]_\alpha^\beta$.

2.4 Vector spaces of linear transformations

Let V and W be two vector spaces of dimensions n and m. Let $\mathcal{L}(V; W)$ denote the set of all linear transformations from V to W, i.e.,

$$\mathcal{L}(V; W) = \{T : V \to W \mid T \text{ is linear}\}.$$

For any two linear transformations S and T in $\mathcal{L}(V; W)$ and $\lambda \in \mathbb{R}$, we define the **sum** $S + T$ and the **scalar multiplication** λS by

$$(S + T)(v) = S(v) + T(v) \quad \text{and} \quad (\lambda S)(v) = \lambda(S(v))$$

for any $v \in V$. Clearly, the sum $S + T$ and the scalar multiplication λS are also linear and satisfy the operational rules of a vector space, so that $\mathcal{L}(V; W)$ becomes a vector space.

Let α and β be two ordered bases for V and W, respectively, and let $T : V \to W$ be a linear transformation. Then the associated matrix $[T]_\alpha^\beta$ of T with respect to these bases is uniquely determined by Theorem 4.9. That is, the function $\phi : \mathcal{L}(V; W) \to M_{m\times n}(\mathbb{R})$ defined by

$$\phi(T) = [T]_\alpha^\beta \in M_{m\times n}(\mathbb{R})$$

for $T \in \mathcal{L}(V; W)$ is well defined (see Section 4.3).

Lemma 2.4.1. *The function $\phi : \mathcal{L}(V; W) \to M_{m\times n}(\mathbb{R})$ is a one-to-one correspondence between $\mathcal{L}(V; W)$ and $M_{m\times n}(\mathbb{R})$.*

Proof. **(1) It is one-to-one:** If $[S]_\alpha^\beta = [T]_\alpha^\beta$ for S and T in $\mathcal{L}(V;W)$, then we have $S = T$ by Corollary 4.4.

(2) It is onto: For any $m \times n$ matrix A (considered as a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$), define a linear transformation $T : V \to W$ by $T = \Psi^{-1} \circ A \circ \Phi$ as the composition of A with the natural isomorphisms $\Phi : V \to \mathbb{R}^n$ and $\Psi : W \to \mathbb{R}^m$. Then clearly $[T]_\alpha^\beta = A$, i.e., ϕ is onto. □

Furthermore, the following lemma shows that ϕ is linear, so that it is in fact an isomorphism from $\mathcal{L}(V;W)$ to $M_{m\times n}(\mathbb{R})$.

Lemma 2.4.2. *Let V and W be vector spaces with ordered bases α and β, respectively, and let $S, T : V \to W$ be linear. Then we have*

$$[S+T]_\alpha^\beta = [S]_\alpha^\beta + [T]_\alpha^\beta \quad \text{and} \quad [\lambda S]_\alpha^\beta = \lambda [S]_\alpha^\beta.$$

Proof. Let $\alpha = \{v_1, \dots, v_n\}$ and $\beta = \{w_1, \dots, w_m\}$. Then we have unique expressions $S(v_j) = \sum_{i=1}^m a_{ij} w_i$ and $T(v_j) = \sum_{i=1}^m b_{ij} w_i$ for each $1 \le j \le n$, so that $[S]_\alpha^\beta = [a_{ij}]$ and $[T]_\alpha^\beta = [b_{ij}]$. Hence

$$(S+T)(v_j) = \sum_{i=1}^m a_{ij} w_i + \sum_{i=1}^m b_{ij} w_i = \sum_{i=1}^m (a_{ij} + b_{ij}) w_i.$$

Thus

$$[S+T]_\alpha^\beta = [S]_\alpha^\beta + [T]_\alpha^\beta.$$

Similarly, $[\lambda S]_\alpha^\beta = \lambda [S]_\alpha^\beta$ (left as an exercise). □

In particular, if $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$, then the vector space $M_{m\times n}(\mathbb{R})$ of $m \times n$ matrices may be identified with the vector space $\mathcal{L}(\mathbb{R}^n;\mathbb{R}^m)$, since such a matrix A is a linear transformation and A itself is the matrix representation of itself with respect to the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$.

The above discussions and results are summarized in the following theorem:

Theorem 2.4.3. *For vector spaces V of dimension n and W of dimension m, the vector space $\mathcal{L}(V;W)$ of all linear transformations from V to W is isomorphic to the vector space $M_{m\times n}(\mathbb{R})$ of all $m \times n$ matrices, and*

$$\dim \mathcal{L}(V;W) = \dim M_{m\times n}(\mathbb{R}) = mn = \dim V \dim W.$$

Proof. Above two lemmas. □

The next theorem shows that the one-to-one correspondence between $\mathcal{L}(V;W)$ and $M_{m\times n}(\mathbb{R})$ preserves not only the vector space structure but also the compositions of linear transformations. Let V, W and Z be vector spaces. Suppose that $S : V \to W$ and $T : W \to Z$ are linear transformations. Then the composition $T \circ S : V \to Z$ is also linear.

Theorem 2.4.4. *Let V, W and Z be vector spaces with ordered bases α, β, and γ, respectively. Suppose that $S : V \to W$ and $T : W \to Z$ are linear transformations. Then*

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

Proof. Let $\alpha = \{v_1, \ldots, v_n\}$, $\beta = \{w_1, \ldots, w_m\}$ and $\gamma = \{z_1, \ldots, z_\ell\}$. Let $[T]_\beta^\gamma = [a_{ij}]$ and $[S]_\alpha^\beta = [b_{pq}]$. Then, for $1 \le i \le n$,

$$\begin{aligned}(T \circ S)(v_i) = T(S(v_i)) &= T\left(\sum_{k=1}^m b_{ki} w_k\right) = \sum_{k=1}^m b_{ki} T(w_k) \\ &= \sum_{k=1}^m b_{ki} \left(\sum_{j=1}^\ell a_{jk} z_j\right) = \sum_{j=1}^\ell \left(\sum_{k=1}^m a_{jk} b_{ki}\right) z_j.\end{aligned}$$

It shows that $[T \circ S]_\alpha^\gamma = [T]_\beta^\gamma [S]_\alpha^\beta$. □

Exercise 2.4.5. Let α be the standard basis for $\mathbb{R}^3$, and let $S, T : \mathbb{R}^3 \to \mathbb{R}^3$ be two linear transformations given by

$$\begin{aligned} S(e_1) &= (2, 2, 1), \quad S(e_2) = (0, 1, 2), \quad S(e_3) = (-1, 2, 1), \\ T(e_1) &= (1, 0, 1), \quad T(e_2) = (0, 1, 1), \quad T(e_3) = (1, 1, 2). \end{aligned}$$

Compute $[S + T]_\alpha$, $[2T - S]_\alpha$ and $[T \circ S]_\alpha$.

Exercise 2.4.6. Let $T : P_2(\mathbb{R}) \to P_2(\mathbb{R})$ be the linear transformation defined by $T(f) = (3+x)f' + 2f$, and let $S : P_2(\mathbb{R}) \to \mathbb{R}^3$ be the one defined by $S(a+bx+cx^2) = (a-b,\ a+b,\ c)$. For a basis $\alpha = \{1, x, x^2\}$ for $P_2(\mathbb{R})$ and the standard basis $\beta = \{e_1, e_2, e_3\}$ for $\mathbb{R}^3$, compute $[S]_\alpha^\beta$, $[T]_\alpha$, and $[S \circ T]_\alpha^\beta$.

Theorem 2.4.7. *Let V and W be vector spaces with ordered bases α and β, respectively, and let $T : V \to W$ be an isomorphism. Then*

$$[T^{-1}]_\beta^\alpha = ([T]_\alpha^\beta)^{-1}.$$

Proof. Since T is invertible, $\dim V = \dim W$, and the matrices $[T]_\alpha^\beta$ and $[T^{-1}]_\beta^\alpha$ are square and of the same size. Thus,

$$[T]_\alpha^\beta [T^{-1}]_\beta^\alpha = [T \circ T^{-1}]_\beta = [id]_\beta$$

is the identity matrix. Hence, $[T^{-1}]_\beta^\alpha = ([T]_\alpha^\beta)^{-1}$. □

In particular, if a linear transformation $T : V \to W$ is an isomorphism, then $[T]_\alpha^\beta$ is an invertible matrix for any bases α for V and β for W.

Exercise 2.4.8. For the vector spaces $P_1(\mathbb{R})$ and $\mathbb{R}^2$, choose the bases $\alpha = \{1, x\}$ for $P_1(\mathbb{R})$ and $\beta = \{e_1, e_2\}$ for $\mathbb{R}^2$, respectively. Let $T : P_1(\mathbb{R}) \to \mathbb{R}^2$ be the linear transformation defined by $T(a + bx) = (a,\ a + b)$.

(1) Show that T is invertible.
(2) Find $[T]_\alpha^\beta$ and $[T^{-1}]_\beta^\alpha$.

2.5 Change of bases

We have seen that, in an n-dimensional vector space V with a fixed basis α, any vector x can be identified with a column vector $[x]_\alpha$ in the n-space $\mathbb{R}^n$ via the natural isomorphism Φ. If another basis β is chosen, then one may get a different column vector $[x]_\beta$. One may ask naturally what is the relation between $[x]_\alpha$ and $[x]_\beta$ for two different bases α and β?

To address this question, let us consider the following example.

Example 2.5.1. Let $V = \mathbb{R}^2$. The coordinate expression of $x = (x, y) \in \mathbb{R}^2$ with respect to the standard basis $\alpha = \{e_1, e_2\}$ is $x = xe_1 + ye_2$, so that $[x]_\alpha = \begin{bmatrix} x \\ y \end{bmatrix}$.

Now let $\beta = \{e_1', e_2'\}$ be another basis for $\mathbb{R}^2$ obtained by rotating α counterclockwise through an angle θ as in the following figure (also see Example 2.1.5), and suppose that the coordinate expression of $x \in \mathbb{R}^2$ with respect to β is written as $x = x'e_1' + y'e_2'$, or $[x]_\beta = \begin{bmatrix} x' \\ y' \end{bmatrix}$.

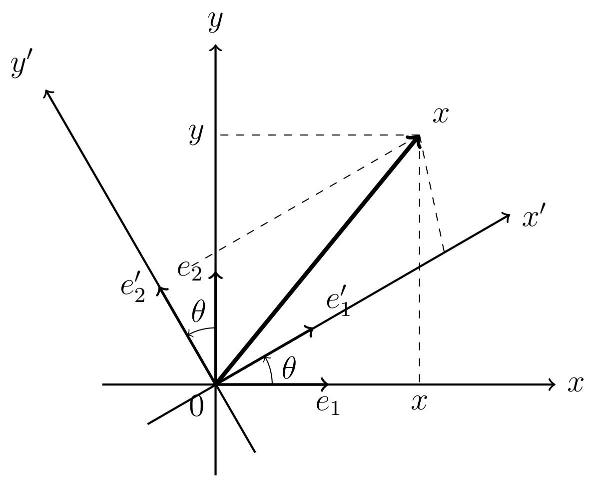

Figure 2.2: Coordinates $\{e_1, e_2\}$ and $\{e_1', e_2'\}$

Then, the expressions of the vectors in β with respect to α are

$$\begin{cases} e_1' = id(e_1') = \cos\theta\, e_1 + \sin\theta\, e_2 \\ e_2' = id(e_2') = -\sin\theta\, e_1 + \cos\theta\, e_2, \end{cases}$$

so

$$[e_1']_\alpha = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}, \quad [e_2']_\alpha = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix}.$$

Therefore, from $x = xe_1 + ye_2$ and

$$x = x'e_1' + y'e_2' = (x'\cos\theta - y'\sin\theta)e_1 + (x'\sin\theta + y'\cos\theta)e_2,$$

one can have the matrix equation:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}, \quad \text{or} \quad [x]_\alpha = [id]_\beta^\alpha [x]_\beta,$$

where

$$[id]_\beta^\alpha = \begin{bmatrix}[e'_1]_\alpha & [e'_2]_\alpha\end{bmatrix} = \begin{bmatrix}\cos\theta & -\sin\theta \\ \sin\theta & \cos\theta\end{bmatrix}.$$

It means that two coordinate vectors $[x]_\alpha$ and $[x]_\beta$ in the 2-space $\mathbb{R}^2$ are related by the associated matrix $[id]_\beta^\alpha$ for the identity transformation id on $\mathbb{R}^2$. Note that

$$[id]_\alpha^\beta = ([id]_\beta^\alpha)^{-1} = \begin{bmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{bmatrix} \quad \text{by the above theorem.} \quad \square$$

In general, if $\alpha = \{v_1, v_2, \ldots, v_n\}$ and $\beta = \{w_1, w_2, \ldots, w_n\}$ are two ordered bases for an n-dimensional vector space V, then any vector $x \in V$ has two expressions:

$$x = \sum_{i=1}^n x_i v_i = \sum_{j=1}^n y_j w_j.$$

In particular, each vector in β is expressed as a linear combination of the vectors in α: say $w_j = id(w_j) = \sum_{i=1}^n q_{ij} v_i$ for $j = 1, 2, \ldots, n$, so that

$$[w_j]_\alpha = [id(w_j)]_\alpha = \begin{bmatrix} q_{1j} \\ \vdots \\ q_{nj}\end{bmatrix}.$$

Then for any $x \in V$,

$$x = \sum_{i=1}^n x_i v_i = \sum_{j=1}^n y_j w_j = \sum_{j=1}^n y_j \sum_{i=1}^n q_{ij} v_i = \sum_{i=1}^n \left(\sum_{j=1}^n q_{ij} y_j\right) v_i.$$

This is equivalent to the matrix equation

$$\boxed{[x]_\alpha = \begin{bmatrix}\sum_{j=1}^n q_{1j} y_j \\ \vdots \\ \sum_{j=1}^n q_{nj} y_j\end{bmatrix} = [id]_\beta^\alpha [x]_\beta,}$$

or

$$\begin{bmatrix}x_1 \\ \vdots \\ x_n\end{bmatrix} = \begin{bmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & \ddots & \vdots \\ q_{n1} & \cdots & q_{nn}\end{bmatrix} \begin{bmatrix}y_1 \\ \vdots \\ y_n\end{bmatrix},$$

where

$$[id]_\beta^\alpha = \begin{bmatrix} q_{11} & \cdots & q_{1n} \\ \vdots & \ddots & \vdots \\ q_{n1} & \cdots & q_{nn}\end{bmatrix} = \begin{bmatrix}[w_1]_\alpha & \cdots & [w_n]_\alpha\end{bmatrix}.$$

This means that any two coordinate vectors of a vector in V with respect to two different ordered bases α and β are related by the matrix representation $[id]_\beta^\alpha$ of the identity transformation on V.

Definition 2.5.2. The matrix representation $[id]_\beta^\alpha$ of the identity transformation $id : V \to V$ with respect to any two ordered bases α and β is called the **basis-change matrix** or the **coordinate-change matrix** from β to α.

Since the identity transformation $id : V \to V$ is invertible, the basis-change matrix $Q = [id]_\beta^\alpha$ is also invertible by the above theorem. If we had taken the expressions of the vectors in the basis α with respect to the basis β: $v_j = id(v_j) = \sum_{i=1}^n p_{ij} w_i$ for $j = 1, 2, \ldots, n$, then $[p_{ij}] = [id]_\alpha^\beta = Q^{-1}$ and

$$[x]_\beta = [id]_\alpha^\beta [x]_\alpha = ([id]_\beta^\alpha)^{-1} [x]_\alpha.$$

Example 2.5.3 (Analytic interpretation of a basis-change matrix)**.** Consider a curve $xy = 1$ on the plane $\mathbb{R}^2$. Find the quadric equation of the curve which is obtained from the curve $xy = 1$ by rotating around the origin clockwise through an angle $\pi/4$.

Solution: Let $\beta = \{e_1', e_2'\}$ be the basis for $\mathbb{R}^2$ obtained by rotating the standard basis $\alpha = \{e_1, e_2\}$ counterclockwise through an angle $\pi/4$, and let $[x]_\alpha = \begin{bmatrix} x \\ y \end{bmatrix}$ and $[x]_\beta = \begin{bmatrix} x' \\ y' \end{bmatrix}$. Then,

$$\begin{bmatrix} x \\ y \end{bmatrix} = [id]_\beta^\alpha \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}.$$

Hence, the equation $xy = 1$ is transformed to

$$1 = xy = \left(\frac{1}{\sqrt{2}}x' - \frac{1}{\sqrt{2}}y'\right)\left(\frac{1}{\sqrt{2}}x' + \frac{1}{\sqrt{2}}y'\right) = \frac{(x')^2}{2} - \frac{(y')^2}{2},$$

which is a hyperbola. (See Figure 4.6.) □

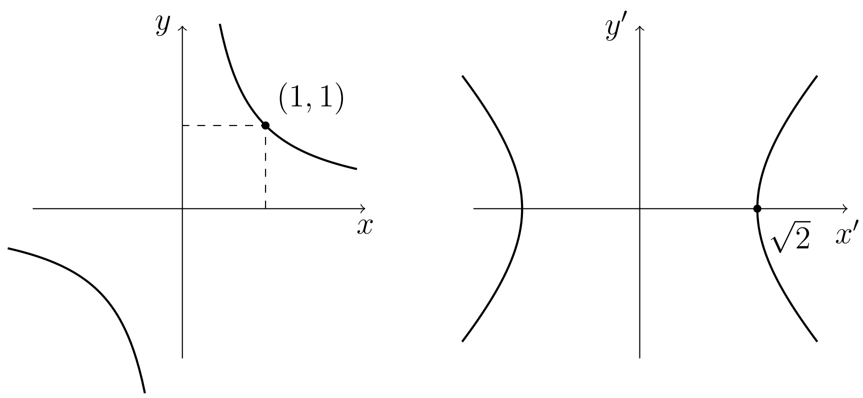

Figure 2.3: The graphs of $xy = 1$ and $\frac{(x')^2}{2} - \frac{(y')^2}{2} = 1$

Example 2.5.4 (Computing a basis-change matrix)**.** Let the 3-space $\mathbb{R}^3$ be equipped with the standard xyz-coordinate system, i.e., with the standard basis $\alpha = \{e_1, e_2, e_3\}$. Take a new $x'y'z'$-coordinate system by rotating the xyz-system around its z-axis counterclockwise through an angle θ, i.e., we take a new basis $\beta = \{e_1', e_2', e_3'\}$ by rotating the basis α about z axis through θ. Then we get

$$[e_1']_\alpha = \begin{bmatrix} \cos\theta \\ \sin\theta \\ 0 \end{bmatrix}, \quad [e_2']_\alpha = \begin{bmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{bmatrix}, \quad [e_3']_\alpha = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Hence, the basis-change matrix from β to α is

$$Q = [id]_\beta^\alpha = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

so

$$[x]_\alpha = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = Q[x]_\beta.$$

Moreover, $Q = [id]_\beta^\alpha$ is invertible and the basis-change matrix from α to β is

$$Q^{-1} = [id]_\alpha^\beta = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

so that

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Exercise 2.5.5. Find the basis-change matrix from a basis α to another basis β for the 3-space $\mathbb{R}^3$, where $\alpha = \{(1,0,1),\ (1,1,0),\ (0,1,1)\}$, $\beta = \{(2,3,1),\ (1,2,0),\ (2,0,3)\}$.

2.6 Similarity

The coordinate expression of a vector in a vector space V depends on the choice of an ordered basis. Hence, the matrix representation of a linear transformation is also dependent on the choice of bases.

Theorem 2.6.1. *Let $T : V \to W$ be a linear transformation from a vector space V with bases α and α' to another vector space W with bases β and β'. Then*

$$[T]_{\alpha'}^{\beta'} = P^{-1}[T]_\alpha^\beta Q,$$

where $Q = [id_V]_{\alpha'}^\alpha$ and $P = [id_W]_{\beta'}^\beta$ are the basis-change matrices.

Proof. Let V and W be two vector spaces of dimensions n and m with two ordered bases α and β, respectively, and let $T : V \to W$ be a linear transformation. Then its associated matrix is $[T]_\alpha^\beta$. If one takes different bases α' and β' for V and W, respectively, then one may get another associated matrix $[T]_{\alpha'}^{\beta'}$ of T. In fact, we have two different expressions

$$\begin{array}{ccccc} [x]_\alpha & \text{and} & [x]_{\alpha'} & \text{in } \mathbb{R}^n & \text{for each } x \in V, \\ [T(x)]_\beta & \text{and} & [T(x)]_{\beta'} & \text{in } \mathbb{R}^m & \text{for } T(x) \in W. \end{array}$$

They are related by the basis-change matrices as follows:

$$[x]_{\alpha'} = [id_V]_\alpha^{\alpha'}[x]_\alpha, \quad \text{and} \quad [T(x)]_{\beta'} = [id_W]_\beta^{\beta'}[T(x)]_\beta.$$

On the other hand, by Theorem 2.3.2, we have

$$[T(x)]_\beta = [T]_\alpha^\beta[x]_\alpha, \quad \text{and} \quad [T(x)]_{\beta'} = [T]_{\alpha'}^{\beta'}[x]_{\alpha'}.$$

Therefore, we get

$$\begin{aligned} [T]_{\alpha'}^{\beta'}[x]_{\alpha'} = [T(x)]_{\beta'} = [id_W]_\beta^{\beta'}[T(x)]_\beta &= [id_W]_\beta^{\beta'}[T]_\alpha^\beta[x]_\alpha \\ &= [id_W]_\beta^{\beta'}[T]_\alpha^\beta[id_V]_{\alpha'}^\alpha[x]_{\alpha'}. \end{aligned}$$

This equation looks messy. However, by Theorem 2.4.4, this relation can be obtained directly from $T = id_W \circ T \circ id_V$ as

$$[T]_{\alpha'}^{\beta'} = [id_W \circ T \circ id_V]_{\alpha'}^{\beta'} = [id_W]_\beta^{\beta'}[T]_\alpha^\beta[id_V]_{\alpha'}^\alpha.$$

Hence,

$$[T]_{\alpha'}^{\beta'} = P^{-1}[T]_\alpha^\beta Q,$$

where $Q = [id_V]_{\alpha'}^\alpha$ and $P = [id_W]_{\beta'}^\beta$ are the basis-change matrices. □

Corollary 2.6.2. *Let $T : V \to V$ be a linear transformation on a vector space V and let α and β be ordered bases for V. Let $Q = [id]_\beta^\alpha$ be the basis-change matrix from β to α. Then*

(1) Q is invertible, and $Q^{-1} = [id]_\alpha^\beta$.

(2) For any $x \in V$, $[x]_\alpha = Q[x]_\beta$.

(3) $[T]_\beta = Q^{-1}[T]_\alpha Q$.

Proof. Taking $W = V$, $\alpha = \beta$ and $\alpha' = \beta'$ in the above theorem, we get $P = Q$ and the the proof follows. □

The relation (3) of $[T]_\beta$ and $[T]_\alpha$ in Corollary 4.16 is called a **similarity**. In general, we have the following definition.

Definition 2.6.3. For any square matrices A and B, A is said to be **similar** to B if there exists a nonsingular matrix Q such that $B = Q^{-1}AQ$.

Note that if A is similar to B, then B is also similar to A. Thus we simply say that A and B are similar. We saw in Corollary 4.16 that if A and B are $n \times n$ matrices representing the same linear transformation T on a vector space V, then A and B are similar.

Example 2.6.4 (Two similar associated matrices)**.** Let $\beta = \{v_1, v_2, v_3\}$ be a basis for the 3-space $\mathbb{R}^3$ consisting of $v_1 = (1,1,0)$, $v_2 = (1,0,1)$ and $v_3 = (0,1,1)$. Let T be the linear transformation on $\mathbb{R}^3$ given by the matrix

$$[T]_\beta = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 3 \\ -1 & 1 & 1 \end{bmatrix}.$$

Let $\alpha = \{e_1, e_2, e_3\}$ be the standard basis. Find the basis-change matrix $[id]_\alpha^\beta$ and $[T]_\alpha$.

Solution: Since $v_1 = e_1 + e_2$, $v_2 = e_1 + e_3$, $v_3 = e_2 + e_3$, we have

$$[id]_\beta^\alpha = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad \text{and} \quad [id]_\alpha^\beta = ([id]_\beta^\alpha)^{-1} = \frac{1}{2}\begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{bmatrix}.$$

Therefore,

$$[T]_\alpha = [id]_\beta^\alpha [T]_\beta [id]_\alpha^\beta = \frac{1}{2}\begin{bmatrix} 4 & 2 & 2 \\ 3 & -1 & 1 \\ -1 & 1 & 7 \end{bmatrix}.$$

Example 2.6.5 (Computing an associated matrix)**.** Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by

$$T(x_1, x_2, x_3) = (2x_1 + x_2,\ x_1 + x_2 + 3x_3,\ -x_2).$$

Let $\alpha = \{e_1, e_2, e_3\}$ be the standard ordered basis for $\mathbb{R}^3$, and let $\beta = \{v_1, v_2, v_3\}$ be another ordered basis consisting of $v_1 = (-1,0,0)$, $v_2 = (2,1,0)$, and $v_3 = (1,1,1)$. Find the associated matrices $[T]_\alpha$ and $[T]_\beta$ for T. Also, show that $T(v_j)$ is the linear combination of the basis vectors in β with the entries of the j-th column of $[T]_\beta$ as its coefficients for $j = 1, 2, 3$.

Solution: One can easily show that

$$[T]_\alpha = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 3 \\ 0 & -1 & 0 \end{bmatrix} \quad \text{and} \quad [id]_\beta^\alpha = \begin{bmatrix} -1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus, with the inverse $[id]_\alpha^\beta = ([id]_\beta^\alpha)^{-1} = \begin{bmatrix} -1 & 2 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$, it follows that

$$[T]_\beta = [id]_\alpha^\beta [T]_\alpha [id]_\beta^\alpha = \begin{bmatrix} 0 & 2 & 8 \\ -1 & 4 & 6 \\ 0 & -1 & -1 \end{bmatrix}.$$

To show the second statement, let $j = 2$. Then $T(v_2) = T(2,1,0) = (5,3,-1)$. On the other hand, the coefficients of $[T(v_2)]_\beta$ are just the entries of the second column of $[T]_\beta$. Therefore,

$$\begin{aligned} T(v_2) &= 2v_1 + 4v_2 - v_3 \\ &= 2(-1,0,0) + 4(2,1,0) - (1,1,1) = (5,3,-1), \end{aligned}$$

as expected. □

The next theorem shows that two similar matrices can be matrix representations of the same linear transformation.

Theorem 2.6.6. *Suppose that an $n \times n$ matrix A represents a linear transformation $T : V \to V$ on a vector space V with respect to an ordered basis $\alpha = \{v_1, v_2, \dots, v_n\}$, i.e., $[T]_\alpha = A$. If $B = Q^{-1}AQ$ for some nonsingular matrix Q, then there exists a basis β for V such that $B = [T]_\beta$ and $Q = [id]_\beta^\alpha$.*

Proof. Let $Q = [q_{ij}]$ and let $w_1, w_2, \dots, w_n$ be the vectors in V defined by

$$\begin{cases} w_1 = q_{11}v_1 + q_{21}v_2 + \cdots + q_{n1}v_n \\ w_2 = q_{12}v_1 + q_{22}v_2 + \cdots + q_{n2}v_n \\ \quad\vdots \\ w_n = q_{1n}v_1 + q_{2n}v_2 + \cdots + q_{nn}v_n. \end{cases}$$

Since Q is nonsingular, the vectors $\beta = \{w_1, w_2, \dots, w_n\}$ form a basis for V. By definition, the basis-change matrix from β to α is $Q = [id]_\beta^\alpha$. Thus, by Corollary 4.16, we have:

$$[T]_\beta = ([id]_\beta^\alpha)^{-1}[T]_\alpha[id]_\beta^\alpha = Q^{-1}AQ = B.$$

□

Then the nonsingularity of $Q = [q_{ij}]$ implies that $\beta = \{v_1, v_2, \dots, v_n\}$ is an ordered basis for V, and Theorem 4.16(3) shows that $[T]_\beta = Q^{-1}[T]_\alpha Q = Q^{-1}AQ = B$ with $Q = [id]_\beta^\alpha$. □

Example 2.6.7 (A matrix similar to an associated matrix is also an associated matrix)**.** Let D be the differential operator on the vector space $P_2(\mathbb{R})$. Given the ordered basis $\alpha = \{1, x, x^2\}$, first note that

$$\begin{aligned} D(1) &= 0 = 0 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ D(x) &= 1 = 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 \\ D(x^2) &= 2x = 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2. \end{aligned}$$

Hence, the matrix representation of D with respect to α is given by

$$[D]_\alpha = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Choose a nonsingular matrix

$$Q = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}, \quad \text{with its inverse} \quad Q^{-1} = \frac{1}{4}\begin{bmatrix} 4 & 0 & 2 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Let

$$B = Q^{-1}[D]_\alpha Q = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix}.$$

Now, we are going to find a basis $\beta = \{v_1, v_2, v_3\}$ so that $B = [D]_\beta$. But, if it is, the matrix Q must be the basis-change matrix $[id]_\beta^\alpha$, and then

$$\begin{aligned} v_1 &= 1 \cdot 1 + 0 \cdot x + 0 \cdot x^2 = 1, \\ v_2 &= 0 \cdot 1 + 2 \cdot x + 0 \cdot x^2 = 2x, \\ v_3 &= -2 \cdot 1 + 0 \cdot x + 4 \cdot x^2 = -2 + 4x^2. \end{aligned}$$

Clearly, one can obtain

$$\begin{aligned} D(1) &= 0 = 0 \cdot 1 + 0 \cdot 2x + 0 \cdot (-2 + 4x^2), \\ D(2x) &= 2 = 2 \cdot 1 + 0 \cdot 2x + 0 \cdot (-2 + 4x^2), \\ D(-2 + 4x^2) &= 8x = 0 \cdot 1 + 4 \cdot 2x + 0 \cdot (-2 + 4x^2), \end{aligned}$$

and

$$[D]_\beta = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{bmatrix},$$

thus, as expected, that $[D]_\beta = B = Q^{-1}[D]_\alpha Q$.

Exercise 2.6.8. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation defined by

$$T(x_1, x_2, x_3) = (x_1 + 2x_2 + x_3,\ -x_2,\ x_1 + 4x_3).$$

Let α be the standard basis, and let $\beta = \{v_1, v_2, v_3\}$ be another ordered basis consisting of $v_1 = (1, 0, 0)$, $v_2 = (1, 1, 0)$, and $v_3 = (1, 1, 1)$ for $\mathbb{R}^3$. Find the associated matrix of T with respect to α and the associated matrix of T with respect to β. Are they similar?

Exercise 2.6.9. Suppose that A and B are similar $n \times n$ matrices. Show that

(1) $\det A = \det B$,
(2) $\operatorname{tr}(A) = \operatorname{tr}(B)$,
(3) $\operatorname{rank} A = \operatorname{rank} B$.

Exercise 2.6.10. Let A and B be $n \times n$ matrices. Show that if A is similar to B, then A^2 is similar to B^2. In general, A^k is similar to B^k for all $k \geq 2$.

2.7 Applications

2.7.1 Dual spaces and adjoint

Definition 2.7.1. Let V be a vector space.

(1) The vector space $\mathcal{L}(V; \mathbb{R})$ of all linear transformations from V to $\mathbb{R}$ is called the **dual space** of V and denoted by V^*.
(2) An element (i.e., a linear transformation) in the dual space $\mathcal{L}(V; \mathbb{R})$ is called a **linear functional** of V.

From the definition, one can say that *any vector space is isomorphic to its dual space*.

Example 2.7.2. The trace function $\operatorname{tr} : M_{n\times n}(\mathbb{R}) \to \mathbb{R}$ is a linear functional of $M_{n\times n}(\mathbb{R})$.

The definite integral of **continuous functions** is one of the most important examples of linear functionals in mathematics.

Example 2.7.3 (Fourier coefficients are linear functionals)**.** Let $C[a,b]$ be the vector space of all continuous real-valued functions on the interval $[a,b]$. The definite integral $\mathcal{I} : C[a,b] \to \mathbb{R}$ defined by

$$\mathcal{I}(f) = \int_a^b f(t)dt$$

is a linear functional of $C[a,b]$. In particular, if the interval is $[0,2\pi]$ and n is an integer, then

$$A_n(f) = \frac{1}{\pi}\int_0^{2\pi} f(t)\cos nt\,dt \quad \text{and} \quad B_n(f) = \frac{1}{\pi}\int_0^{2\pi} f(t)\sin nt\,dt$$

are linear functionals, called the n-th **Fourier coefficients** of f. □

For a matrix A regarded as a linear transformation $A : \mathbb{R}^n \to \mathbb{R}^m$, the transpose A^T of A is another linear transformation $A^T : \mathbb{R}^m \to \mathbb{R}^n$. For a linear transformation $T : V \to W$ from a vector space V to W, one can naturally ask what its *transpose* is and what the definition is. In this section, we discuss this problem.

Recall that a linear functional $T : V \to \mathbb{R}$ is completely determined by the values on a basis for V. Let $\alpha = \{v_1, v_2, \dots, v_n\}$ be a basis for a vector space V. For each $i = 1, 2, \dots, n$, define a linear functional

$$v_i^* : V \to \mathbb{R}$$

by $v_i^*(v_j) = \delta_{ij}$ for each $j = 1, 2, \dots, n$. Then, for any $x = \sum a_i v_i \in V$, we have $v_i^*(x) = a_i$, which is the i-th coordinate of x with respect to α. Thus, the functional v_i^* is called the i-th **coordinate function** with respect to the basis α.

Theorem 2.7.4. *The set $\alpha^* = \{v_1^*, v_2^*, \dots, v_n^*\}$ of coordinate functions forms a basis for the dual space V^*, and for any $T \in V^*$ we have*

$$T = \sum_{i=1}^{n} T(v_i)v_i^*.$$

Definition 2.7.5. For a basis $\alpha = \{v_1, v_2, \dots, v_n\}$ for a vector space V, the basis $\alpha^* = \{v_1^*, v_2^*, \dots, v_n^*\}$ for V^* is called the **dual basis** of α.

Example 2.7.6 (Computing a dual basis)**.** Let $\alpha = \{v_1, v_2\}$ be a basis for $\mathbb{R}^2$, where $v_1 = (1,2)$ and $v_2 = (1,3)$. To find its dual basis $\alpha^* = \{v_1^*, v_2^*\}$ of α, we consider the equations

$$\begin{aligned} 1 &= v_1^*(v_1) = v_1^*(e_1) + 2v_1^*(e_2), \\ 0 &= v_1^*(v_2) = v_1^*(e_1) + 3v_1^*(e_2). \end{aligned}$$

Solving these equations, we obtain that $v_1^*(e_1) = 3$ and $v_1^*(e_2) = -1$. Thus $v_1^*(x,y) = 3x - y$. Similarly, it can be shown that $v_2^*(x,y) = -2x + y$. □

Exercise 2.7.7. Let $\alpha = \{(1,0,1),\ (1,2,1),\ (0,0,1)\}$ be a basis for $\mathbb{R}^3$. Find the dual basis α^*.

For a given linear transformation $T : V \to W$, one can define $T^* : W^* \to V^*$ by $T^*(g) = g \circ T$ for any $g \in W^*$. In fact, for any linear functional $g \in W^*$, i.e., $g : W \to \mathbb{R}$, the composition $g \circ T : V \to \mathbb{R}$ given by $g \circ T(x) = g(T(x))$ for $x \in V$ defines a linear functional on V, i.e., $T^*(g) = g \circ T \in V^*$.

Lemma 2.7.8. *The transformation $T^* : W^* \to V^*$ defined by $T^*(g) = g \circ T$ for $g \in W^*$ is a linear transformation. It is called the **adjoint** (or **transpose**) of T.* □

Proof. For any $f, g \in W^*$, $a, b \in \mathbb{R}$ and $x \in V$,

$$T^*(af + bg)(x) = af(T(x)) + bg(T(x)) = (aT^*(f) + bT^*(g))(x).$$

□

Example 2.7.9 ($(id_V)^* = id_{V^*}$ and $(T \circ S)^* = S^* \circ T^*$)**.** (1) Let $id : V \to V$ be the identity transformation on a vector space V. Then for any $g \in V^*$, $id^*(g) = g \circ id = g$. Hence, the adjoint $id^* : V^* \to V^*$ is the identity transformation on V^*, i.e., $id^* = id$.

(2) Let $S : U \to V$ and $T : V \to W$ be two linear transformations. Then for any $g \in W^*$, we have

$$\begin{aligned}(T \circ S)^*(g) &= g \circ (T \circ S) = (g \circ T) \circ S \\ &= T^*(g) \circ S = S^*(T^*(g)) = (S^* \circ T^*)(g).\end{aligned}$$

It shows that $(T \circ S)^* = S^* \circ T^*$. □

Now, if $S : V \to W$ is an isomorphism, then $(S^{-1})^* \circ S^* = (S \circ S^{-1})^* = id^* = id$ shows that $S^* : W^* \to V^*$ is also an isomorphism.