

Section 4. Linear Transformation and Its Matrix Expression





1 The concept of linear transformation

1. Mapping

Definition 1. Let A, B be two non - empty sets. To each α in A , by some rule T , corresponds one and only one definite β in B . Then the rule T is called a mapping from A to B , which is denoted by $\beta = T(\alpha)$ or $\beta = T\alpha$, ($\alpha \in A$). β is called the image of α and α is called the pre - image of β .





For $A_0 \subseteq A$, the set of the images of all elements in A_0 is denoted by $T(A)$, i.e.

$$T(A_0) = \{\beta = T(\alpha) \mid \alpha \in A_0\},$$

Apparently, $T(A) \subset B$.

Clearly, mapping is an extension of the concept of function.





Definition 2. Let V be a linear space T a mapping from V to V . If T satisfies the following conditions, then T is called a linear transformation.

$$(1) \quad \forall \alpha, \beta \in V, T(\alpha + \beta) = T(\alpha) + T(\beta).$$

$$(2) \quad \forall \lambda \in \mathbf{R}, \alpha \in V, T(\lambda\alpha) = \lambda T(\alpha).$$

Example 1. In $P[x]_3$, the differentiation operation is a linear transformation.

Proof. $\forall p = a_3 x^3 + a_2 x^2 + a_1 x + a_0 \in P[x]_3,$

$$Dp = 3a_3 x^2 + 2a_2 x + a_1,$$





$$q = b_3 x^3 + b_2 x^2 + b_1 x + b_0 \in P[x]_3,$$

$$Dq = 3b_3 x^2 + 2b_2 x + b_1,$$

Hence, $D(p + q)$

$$= D[(a_3 + b_3)x^3 + (a_2 + b_2)x^2 + (a_1 + b_1)x + (a_0 + b_0)]$$

$$= 3(a_3 + b_3)x^2 + 2(a_2 + b_2)x + (a_1 + b_1)$$

$$= (3a_3 x^2 + 2a_2 x + a_1) + (3b_3 x^2 + 2b_2 x + b_1)$$

$$= Dp + Dq;$$





$$\begin{aligned}D(kp) &= D(k a_3 x^3 + k a_2 x^2 + k a_1 x + k a_0) \\&= k(3 a_3 x^2 + 2 a_2 x + a_1) \\&= kDp.\end{aligned}$$

Hence, T is a linear transformation.

Remark 1. If $T(p) = a_0$, then T is also a linear transformation.

The reason : $T(p + q) = a_0 + b_0 = T(P) + T(q)$.

$$T(\lambda p) = \lambda a_0 = \lambda T(p).$$





Remark 2. If $T(p) = 1$, then T is not a linear transformation.

The reason : $T(p + q) = 1$,

but $T(p) + T(q) = 1 + 1 = 2$,

As a result, $T(p + q) \neq T(p) + T(q)$.





Example 2. Let T be defined by

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Then T is a linear transformation over xOy plane.

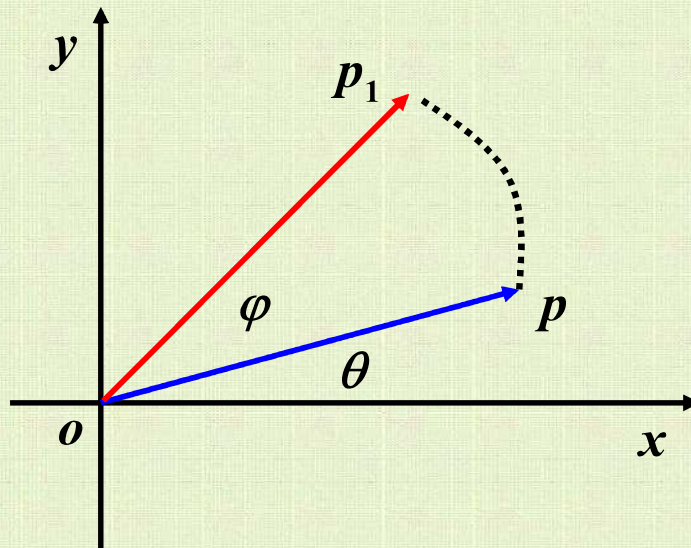
Solution. Let $\begin{cases} x = r \cos \theta, \\ y = r \sin \theta, \end{cases}$ so

$$\begin{aligned} T \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} x \cos \varphi - y \sin \varphi \\ x \sin \varphi + y \cos \varphi \end{pmatrix} \\ &= \begin{pmatrix} r \cos \theta \cos \varphi - r \sin \theta \sin \varphi \\ r \cos \theta \sin \varphi + r \sin \theta \cos \varphi \end{pmatrix} = \begin{pmatrix} r \cos(\theta + \varphi) \\ r \sin(\theta + \varphi) \end{pmatrix}, \end{aligned}$$





Geometric explanation : T rotates a vector by an angle of φ anti-clockwise .





Example 3. The set of all continuous functions on the interval $[a,b]$ is a linear space. In this space, define

$$T(f(x)) = \int_a^x f(t)dt$$

Then T is a linear transformation.

Proof. Let $f(x) \in V, g(x) \in V$.

$$\begin{aligned} \text{Then } T[f(x) + g(x)] &= \int_a^x f(t) + g(t)dt \\ &= \int_a^x f(t)dt + \int_a^x g(t)dt \\ &= T[f(x)] + T[g(x)] \end{aligned}$$





$$T(kf(x)) = \int_a^x kf(t)dt = k \int_a^x f(t)dt = kT[f(x)].$$

Example 4. The following mapping in any linear space V is a linear transformation.

$$E(\alpha) = \alpha, \quad \alpha \in V.$$

Proof. For any $\alpha, \beta \in V$,

$$E(\alpha + \beta) = \alpha + \beta = E(\alpha) + E(\beta)$$

$$E(k\alpha) = k\alpha = kE(\alpha).$$

Hence E is a linear transformation, which is called identity transformation.





Example 5. The zero transformation in any linear space V defined by $O(\alpha) = \mathbf{0} (\forall \alpha \in V)$ is a linear transformation.

Proof. Let $\alpha, \beta \in V$. Then

$$\mathbf{0}(\alpha + \beta) = \mathbf{0} = \mathbf{0} + \mathbf{0} = \mathbf{0}(\alpha) + \mathbf{0}(\beta)$$

$$\mathbf{0}(k\alpha) = \mathbf{0} = k\mathbf{0} = k\mathbf{0}(\alpha).$$

So the zero transformation is linear.





Example 6. In R^3 , define

$$T(x_1, x_2, x_3) = (x_1^2, x_2 + x_3, 0)$$

Then T is not a linear transformation.

Proof. $\forall \alpha = (a_1, a_2, a_3), \beta = (b_1, b_2, b_3) \in R^3$,

$$\begin{aligned} T(\alpha + \beta) &= T(a_1 + b_1, a_2 + b_2, a_3 + b_3) \\ &= ((a_1 + b_1)^2, a_2 + a_3 + b_2 + b_3, 0) \\ &\neq (a_1^2, a_2 + a_3, 0) + (b_1^2, b_2 + b_3, 0) \\ &= T(\alpha) + T(\beta). \end{aligned}$$





2. Properties of linear transformations

1. $T(0) = 0$, $T(-\alpha) = -T(\alpha)$;

2. If $\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m$, then

$$T\beta = k_1T\alpha_1 + k_2T\alpha_2 + \cdots + k_mT\alpha_m;$$

3. If $\alpha_1, \alpha_2, \dots, \alpha_m$ is linearly dependent, then

$T\alpha_1, T\alpha_2, \dots, T\alpha_m$ is also linearly dependent.

Note. Linear independence of $\alpha_1, \alpha_2, \dots, \alpha_m$ does not necessarily imply linear independence of $T\alpha_1, T\alpha_2, \dots, T\alpha_m$.





4. Let T be a linear transformation of V_n . Then the set $T(V_n)$ of the images of V_n is a subspace of V_n .

Proof. Let $\beta_1, \beta_2 \in T(V_n)$, so there exist $\alpha_1, \alpha_2 \in V_n$ such that $T\alpha_1 = \beta_1, T\alpha_2 = \beta_2$. Thus

$$\beta_1 + \beta_2 = T\alpha_1 + T\alpha_2 = T(\alpha_1 + \alpha_2) \in T(V_n).$$

$$k\beta_1 = kT\alpha_1 = T(k\alpha_1) \in T(V_n).$$

Since $T(V_n) \subset V_n$, it is a subspace of V_n .





5. Let $S_T = \{\alpha \mid \alpha \in V_n, T\alpha = \mathbf{0}\}$. Then S_T is a subspace of V_n which is called the kernel of T .

Proof. If $\alpha_1, \alpha_2 \in S_T$, then $T\alpha_1 = \mathbf{0}, T\alpha_2 = \mathbf{0}$, and thus $T(\alpha_1 + \alpha_2) = T\alpha_1 + T\alpha_2 = \mathbf{0}$ i.e. $\alpha_1 + \alpha_2 \in S_T$.

If $\alpha_1 \in S_T, k \in R$, then

$$T(k\alpha_1) = kT\alpha_1 = k\mathbf{0} = \mathbf{0} \text{ i.e. } k\alpha_1 \in S_T.$$

As a result, S_T is closed for the linear operations .

From $S_T \subset V_n$, it follows that S_T is a subspace of V_n .





Example 7. Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = (\alpha_1, \alpha_2, \cdots, \alpha_n),$$

where $\alpha_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}$. Define a mapping T in R^n by

$$T(x) = Ax (x \in R^n).$$

Then T is a linear transformation .





Indeed , if $a, b \in R^n$, then

$$T(a + b) = A(a + b) = Aa + Ab = T(a) + T(b);$$

$$T(ka) = A(ka) = kAa = kT(a).$$

Clearly,

$$T(R^n) = \{y = x_1\alpha_1 + x_2\alpha_2 + \cdots + x_n\alpha_n \mid x_1, x_2, \cdots, x_n \in R\}$$

whereas the kernel S_T of T is the solution space of the system of linear equations $Ax = 0$.





3. The matrix expression of a linear transformation

As we know, if

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = (\alpha_1, \alpha_2, \cdots, \alpha_n),$$

where $\alpha_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}$, define T by





$T(x) = Ax(x \in R^n)$, then T is a linear transformation.

Assume e_1, e_2, \dots, e_n are the unit coordinate vectors.

$$\text{Then } Ae_1 = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \alpha_1, \dots,$$

$$Ae_n = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} = \alpha_n,$$





$$\text{i.e. } \alpha_i = Ae_i = T(e_i) \quad (i = 1, 2, \dots, n)$$

Conclusion : If $T(x) = Ax$, then the columns of A are $T(e_i)$ ($i = 1, 2, \dots, n$).

Conversely, if a linear transformation T such that $T(e_i) = \alpha_i$ ($i = 1, 2, \dots, n$), then

$$\begin{aligned} T(x) &= T[(e_1, e_2, \dots, e_n)x] \\ &= T(x_1e_1 + x_2e_2 + \dots + x_n e_n) \\ &= x_1T(e_1) + x_2T(e_2) + \dots + x_nT(e_n) \\ &= (T(e_1), T(e_2), \dots, T(e_n))x \\ &= (\alpha_1, \alpha_2, \dots, \alpha_n)x = Ax. \end{aligned}$$





In other words,

Each linear transformation T of R^n can be expressed by

$$T(x) = Ax \quad (x \in R^n)$$

where $A = (T(e_1), T(e_2), \dots, T(e_n))$

$$= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

with e_1, e_2, \dots, e_n being the unit coordinate vectors.





denoted by $T(\alpha_1, \alpha_2, \dots, \alpha_n) = (T(\alpha_1), T(\alpha_2), \dots, T(\alpha_n))$,

then we have

$$T(\alpha_1, \alpha_2, \dots, \alpha_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)A$$

where $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$

We shall call A the matrix of the linear transformation under the given basis $\alpha_1, \alpha_2, \dots, \alpha_n$





Obviously, A is uniquely determined by $T(\alpha_1), \dots, T(\alpha_n)$.

Now assume A is the matrix of T under the given basis $\alpha_1, \alpha_2, \dots, \alpha_n$, i.e.

$$T(\alpha_1, \alpha_2, \dots, \alpha_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)A.$$

Let us check what conditions should be satisfied by T .





$\forall \alpha \in V_n$, let $\alpha = \sum_{i=1}^n x_i \alpha_i$, then

$$T(\alpha) = T\left(\sum_{i=1}^n x_i \alpha_i\right) = \sum_{i=1}^n x_i T(\alpha_i)$$

$$= (T(\alpha_1), T(\alpha_2), \dots, T(\alpha_n)) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

$$= (\alpha_1, \alpha_2, \dots, \alpha_n) A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix},$$





namely,

$$T \left[(\alpha_1, \alpha_2, \dots, \alpha_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right] = (\alpha_1, \alpha_2, \dots, \alpha_n) A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

The above expression uniquely determines the linear transformation T , whose matrix is A .





Conclusion:

When a basis in a linear space is selected, a linear transformation T uniquely determines a matrix A .

Conversely, a matrix A uniquely determines a linear transformation T .

In other words, linear transformation and matrix are uniquely determined by each other after a basis is selected.





From

$$T \left[(\alpha_1, \alpha_2, \dots, \alpha_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right] = (\alpha_1, \alpha_2, \dots, \alpha_n) A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

It follows that under $\alpha_1, \alpha_2, \dots, \alpha_n$,

the coordinate of α is $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$.





The coordinate of $T(\alpha)$ is $A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$,

or equivalently,

$$T(\alpha) = A\alpha.$$





Example 1. In $P[x]_4$, take the basis

$$p_1 = x^3, p_2 = x^2, p_3 = x, p_4 = 1.$$

Find the matrix of the differentiation operation D .

Solution.

$$\begin{cases} D p_1 = 3x^2 = 0 p_1 + 3 p_2 + 0 p_3 + 0 p_4, \\ D p_2 = 2x = 0 p_1 + 0 p_2 + 2 p_3 + 0 p_4, \\ D p_3 = 1 = 0 p_1 + 0 p_2 + 0 p_3 + 1 p_4, \\ D p_4 = 0 = 0 p_1 + 0 p_2 + 0 p_3 + 0 p_4, \end{cases}$$





Hence the matrix of D under the basis in question is

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$





Example 2.

In the linear space $R[x]_n$, define the mapping

$$\sigma(f(x)) = \frac{d}{dx} f(x), f(x) \in R[x]_n.$$

By properties of derivatives, σ is a linear transformation of $R[x]_n$, which is called differential transformation.

Take the basis $1, x, x^2, \dots, x^{n-1}$ in $R[x]_n$.





Then we have

$$\sigma(1) = 0, \quad \sigma(x) = 1, \quad \sigma(x^2) = 2x,$$
$$\dots\dots\dots, \quad \sigma(x^{n-1}) = (n-1)x^{n-2}$$

So the matrix of σ under $1, x, x^2, \dots, x^{n-1}$ is

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & n-1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$





Example 3. In R^3 , let T denote the projection on the xOy plane, i.e.

$$T(x\vec{i} + y\vec{j} + z\vec{k}) = x\vec{i} + y\vec{j},$$

(1) Find the matrix of T under the basis $\vec{i}, \vec{j}, \vec{k}$.

(2) Find the matrix of T under the basis $\alpha = \vec{i}, \beta = \vec{j}, \gamma = \vec{i} + \vec{j} + \vec{k}$.

Solution.

(1)

$$\begin{cases} T\vec{i} = \vec{i}, \\ T\vec{j} = \vec{j}, \\ T\vec{k} = \vec{0}, \end{cases}$$

$$\text{i.e. } T(\vec{i}, \vec{j}, \vec{k}) = (\vec{i}, \vec{j}, \vec{k}) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$





$$(2) \quad \begin{cases} T\alpha = \vec{i} = \alpha, \\ T\beta = \vec{j} = \beta, \\ T\gamma = \vec{i} + \vec{j} = \alpha + \beta, \end{cases}$$

$$\text{i.e. } T(\alpha, \beta, \gamma) = (\alpha, \beta, \gamma) \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

This example shows that a linear transformation has different matrix under different bases.

However, there exists some relationship among these matrices, which is explicitly shown in the following theorem.





5. The matrices of a linear transformation under different bases

Theorem 1. Let $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ and $\{\beta_1, \beta_2, \dots, \beta_n\}$ be two bases in the linear space V_n . If the transition matrix from $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ to $\{\beta_1, \beta_2, \dots, \beta_n\}$ is P , and if the matrices of a linear transformation T under $\{\alpha_1, \alpha_2, \dots, \alpha_n\}$ and $\{\beta_1, \beta_2, \dots, \beta_n\}$ are A and B respectively, then we have

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





Proof. Since $(\beta_1, \beta_2, \dots, \beta_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)P$

$$T(\alpha_1, \alpha_2, \dots, \alpha_n) = (\alpha_1, \alpha_2, \dots, \alpha_n)A,$$

$$T(\beta_1, \beta_2, \dots, \beta_n) = (\beta_1, \beta_2, \dots, \beta_n)B$$

Thus

$$\begin{aligned}(\beta_1, \beta_2, \dots, \beta_n)B &= T(\beta_1, \beta_2, \dots, \beta_n) \\ &= T[(\alpha_1, \alpha_2, \dots, \alpha_n)P] \\ &= T[(\alpha_1, \alpha_2, \dots, \alpha_n)]P\end{aligned}$$





$$= (\alpha_1, \alpha_2, \dots, \alpha_n)AP$$

$$= (\beta_1, \beta_2, \dots, \beta_n)P^{-1}AP$$

Since $\beta_1, \beta_2, \dots, \beta_n$ are linearly independent,

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

It follows from Theorem 1 that the matrices of a linear transformation under two bases are similar.





Example 4.

Let the matrix of a linear transformation T of V_2 under $\{\alpha_1, \alpha_2\}$ is

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

Find the matrix of T under the basis $\{\alpha_2, \alpha_1\}$.

Solution. $(\alpha_2, \alpha_1) = (\alpha_1, \alpha_2) \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$

i.e. $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$ and thus $P^{-1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$





Consequently, the matrix of T under the basis $\{\alpha_2, \alpha_1\}$ is

$$\begin{aligned} B &= \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} \\ &= \begin{pmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{1} & \mathbf{0} \end{pmatrix} \\ &= \begin{pmatrix} a_{22} & a_{21} \\ a_{12} & a_{11} \end{pmatrix}. \end{aligned}$$





Definition 2. The dimension of the image space $T(V_n)$ of a linear transformation T is called the rank of T .

Conclusion :

(1) Let A be the matrix of T under a basis. Then

The rank of A equals the rank of T .

(2) If the rank of T is r , then the dimension of the kernel S_T of T is $n - r$.





Example 5. Let the matrix of the linear transformation σ of 3-dimensional space V under the basis $\{\alpha_1, \alpha_2, \alpha_3\}$ be

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

Find the matrix of σ under the basis $\{\alpha_2, \alpha_3, \alpha_1\}$.

Solution. By the given condition,

$$\sigma(\alpha_1, \alpha_2, \alpha_3) = (\alpha_1, \alpha_2, \alpha_3) \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$





$$\text{i.e.} \quad \begin{cases} \sigma(\alpha_1) = \alpha_1 + 4\alpha_2 + 7\alpha_3 \\ \sigma(\alpha_2) = 2\alpha_1 + 5\alpha_2 + 8\alpha_3 \\ \sigma(\alpha_3) = 3\alpha_1 + 6\alpha_2 + 9\alpha_3 \end{cases}$$

$$\text{and hence} \quad \begin{cases} \sigma(\alpha_2) = 5\alpha_2 + 8\alpha_3 + 2\alpha_1 \\ \sigma(\alpha_3) = 6\alpha_2 + 9\alpha_3 + 3\alpha_1 \\ \sigma(\alpha_1) = 4\alpha_2 + 7\alpha_3 + \alpha_1 \end{cases}$$

Therefore, the matrix of σ under $\{\alpha_2, \alpha_3, \alpha_1\}$ is

$$B = \begin{pmatrix} 5 & 8 & 2 \\ 6 & 9 & 3 \\ 4 & 7 & 1 \end{pmatrix}.$$

