UNIT 9

LINEAR TRANSFORMATIONS AND MATRICES

Structure **Page No.**

9.1 Introduction
Objectives
9.2 Matrix of a Linear Transformation
Sum and Multiplication by Scalars
Dimension of $\mathbf{M_{n\times n}(F)}$ over $\mathbf{F}$
9.3 Some Types of Matrices
Diagonal Matrix
Triangular Matrix
9.4 Matrix Multiplication
Matrix of the Composition of Linear Transformations
9.5 Invertible Matrices
Inverse of Matrix
Matrix of Change of Basis
9.6 Summary
9.7 Solutions / Answers

9.1 INTRODUCTION

In Unit 1, we discussed matrices. In Unit 2, we saw that $M_{n,\,m}(F)$ is a vector space over F. In this unit, we will look at the matrices; the matrices will be representations of linear transformations. Such a representation will enable us to compute several things like the rank and nullity of a linear transformation. We will also see that composition of linear transformations translates to multiplication of matrices. This gives us an insight into matrix multiplication which looks very contrived otherwise.

Matrices are intimately connected with linear transformations. In this unit we will bring out this link. We will first derive algebraic operations on matrices from the corresponding operations on linear transformations. In Block 4 we will often refer to the material on change of bases, so do spend some time on Sec.7.6.

To realise the deep connection between matrices and linear transformations, you should go back to the exact spot in Unit 7 and 8 to which frequent references are made.

Objectives

After studying this unit, you should be able to:

- define and give examples of various types of matrices;
- define a linear transformation, if you know its associated matrix;
- evaluate the sum, difference, product and scalar multiple of matrices;
- obtain the transpose and conjugate of a matrix;
- determine if a given matrix is invertible;
- obtain the inverse of a matrix;
- discuss the effect that the change of basis has on the matrix of a linear transformation.

9.2 MATRIX OF A LINEAR TRANSFORMATION

We will now obtain a matrix that corresponds to a given linear transformation. You will see how easy it is to go from matrices to linear transformations, and back.

Let U and V be vector spaces over a field F, of dimensions n and m, respectively. Let $B_1 = \{e_1, \ldots, e_n\}$ be an ordered basis of U, and $B_2 = \{f_1, \ldots, f_m\}$ be an ordered basis of V. (By an **ordered basis** we mean that the order in which the elements of the basis are written is fixed. Thus, an ordered basis $\{e_1, e_2\}$ is not equal to an ordered basis $\{e_2, e_1\}$.)

Given a linear transformation $T : U \to V$, we will associate a matrix to it. For this, we consider $T(e_1), \ldots, T(e_n)$, which are all elements of V and hence, they are linear combinations of $f_1, \ldots, f_m$. Thus, there exist mn scalars α_{ij}, such that

$$\begin{aligned} T(e_1) &= \alpha_{11}f_1 + \alpha_{21}f_2 + \cdots + \alpha_{m1}f_m \\ &\vdots \\ T(e_j) &= \alpha_{1j}f_1 + \alpha_{2j}f_2 + \cdots + \alpha_{mj}f_m \\ &\vdots \\ T(e_n) &= \alpha_{1n}f_1 + \alpha_{2n}f_2 + \cdots + \alpha_{nm}f_m \end{aligned}$$

From these n equations we form an $m \times n$ matrix whose first row consists of the coefficients of the first equation, second column consists of the coefficients of the second equation, and so on. This matrix

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix}$$

is called the **matrix of $\mathbf{T}$ with respect to the bases $\mathbf{B_1}$ and $\mathbf{B_2}$**. Notice that **the coordinate vector of $\mathbf{T(e_j)}$ is the $\mathbf{jth}$ column of $\mathbf{A}$**.

We use the notation $[T]_{B_2}^{B_1}$ for this matrix. Thus, to obtain $[T]_{B_2}^{B_1}$ we consider $T(e_i) \ \forall \ e_i \in B_1,$ and write them as linear combination of the elements of B_2.

If $T \in L(V, V), B$ is basis of V and we take $B_1 = B_2 = B,$ we call $[T]_B^B$ the matrix of T with respect to the basis $B,$ and we write this as $[T]_B$.

Remark 1: Why do we insist on ordered bases? What happens if we interchange the order of the elements in B_1 to $\{e_n, e_1, \ldots, e_{n-1}\}$? The matrix $[T]_{B_2}^{B_1}$ also changes, the last column becoming the first column now. Similarly, if we change the positions of the $f_i's$ in $B_2,$ the rows of $[T]_{B_2}^{B_1}$ will get interchanged.

Thus, to obtain a unique matrix corresponding to $T,$ we must insist on B_1 and B_2 being ordered bases. **Henceforth, while discussing the matrix of a linear mapping, we will always assume that our bases are ordered bases.**

We will now give an example, followed by some exercises.

Example 1: Consider the linear operator $T : \mathbf{R}^3 \to \mathbf{R}^2 : T(x, y, z) = (x, y).$ Choose bases B_1 and B_2 of $\mathbf{R}^3$ and $\mathbf{R}^2,$ respectively. Then obtain $[T]_{B_2}^{B_1}.$

Solution: Let $B_1 = \{e_1, e_2, e_3\},$ where $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1).$ Let $B_2 = \{f_1, f_2\},$ where $f_1 = (1, 0), f_2 = (0, 1).$ Note that B_1 and B_2 are the standard bases of $\mathbf{R}^3$ and $\mathbf{R}^2,$ respectively.

$$T(e_1) = (1, 0) = f_1 = 1.f_1 + 0.f_2$$
$$T(e_2) = (0, 1) = f_2 = 0.f_1 + 1.f_2$$
$$T(e_3) = (0, 0) = 0f_1 + 0f_2.$$

Thus, $[T]_{B_2}^{B_1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$

***

E1) Choose two other bases B_1' and B_2' of $\mathbf{R}^3$ and $\mathbf{R}^2,$ respectively. (In Unit 4 you came across a lot of bases of both these vector spaces.) for T in the example above, give the matrix $[T]_{B_2'}^{B_1'}$

What E1 shows us is that the matrix of a transformation depends on the bases that we use for obtaining it. The next two exercises also bring out the same fact.

E2) Write the matrix of the linear transformation $T : \mathbf{R}^3 \to \mathbf{R}^2 : T(x, y, z) = (x + 2y + 2z, 2x + 3y + 4z)$ with respect to the standard bases of $\mathbf{R}^3$ and $\mathbf{R}^2$.

E3) What is the matrix of T, in E2, with respect to the bases

$B_1' = \{(1, 0, 0), (0, 1, 0), (1, -2, 1)\}$ and

$B_2' = \{(1, 2), (2, 3)\}$?

E4) Let V be the vector space of polynomials over $\mathbf{R}$ of degree ≤ 3, in the variable t. Let $D : V \to V$ be the differential operator given in Unit 5 (E6, when $n = 3$). Show that the matrix of D with respect to the basis $\{1, t, t^2, t^3\}$ is

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

So far, given a linear transformation, we have obtained a matrix from it. This works the other way also. That is, given a matrix we can define a linear transformation corresponding to it.

Example 2: Describe $T : \mathbf{R}^3 \to \mathbf{R}^3$ such that $[T]_B = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}$, when B is the standard basis of $\mathbf{R}^3$.

Solution: Let $B = \{e_1, e_2, e_3\}$. Now, we are given that

$$T(e_1) = 1.e_1 + 2.e_2 + 3.e_3$$
$$T(e_2) = 2.e_1 + 3.e_2 + 1.e_3$$
$$T(e_3) = 4.e_1 + 1.e_2 + 2.e_3$$

You know that any element of $\mathbf{R}^3$ is $(x, y, z) = xe_1 + ye_2 + ze_3$.

Therefore, $T(x, y, z) = T(xe_1 + ye_2 + ze_3)$

$= xT(e_1) + yT(e_2) + zT(e_3)$, since T is linear.

$= x(e_1 + 2e_2 + 3e_3) + y(2e_1 + 3e_2 + e_3) + z(4e_1 + e_2 + 2e_3)$

$= (x + 2y + 4z)e_1 + (2x + 3y + z)e_2 + (3x + y + 2z)e_3$

$= (x + 2y + 4z, 2x + 3y + z, 3x + y + 2z)$

$\therefore T : \mathbf{R}^3 \to \mathbf{R}^3$ is defined by

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

***

Try the following exercises now.

E5) Describe $T: \mathbf{R}^3 \to \mathbf{R}^2$ such that $[T]_{B_2}^{B_1} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$, where B_1 and B_2 are the standard bases of $\mathbf{R}^3$ and $\mathbf{R}^2$, respectively.

E6) Find the linear operator $T: \mathbf{C} \to \mathbf{C}$ whose matrix, with respect to the basis $\{1, i\}$ is $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. (Note that $\mathbf{C}$, the field of complex numbers, is a vector space over $\mathbf{R}$, of dimension 2.)

Now we are in a position to define the sum of matrices and multiplication of a matrix by a scalar.

9.2.1 Sum and Multiplication by Scalars

In Unit 7 you studied about the sum and scalar multiples of linear transformation. In the following theorem we will see what happens to the matrices associated with the linear transformations that are sums or scalar multiples of given linear transformations.

Theorem 1: Let U and V be vectors spaces over $\mathbf{F}$, of dimensions n and m, respectively. Let B_1 and B_2 be arbitrary bases of U and V, respectively. (Let us abbreviate $[T]_{B_2}^{B_1}$ to $[T]$ during this theorem.) Let $S, T \in L(U, V)$ and $\alpha \in \mathbf{F}$. Suppose $[S] = [a_{ij}], [T] = [b_{ij}]$. Then

$[S+T] = [a_{ij} + b_{ij}]$, and

$[\alpha S] = [\alpha a_{ij}]$.

Proof: Suppose $B_1 = \{e_1, e_2, \ldots, e_n\}$ and $B_2 = \{f_1, f_2, \ldots, f_m\}$. Then all the matrices to be considered here will be of size $m \times n$.
Now, by our hypothesis,

$$S(e_j) = \sum_{i=1}^{m} a_{ij} f_i \ \forall \ j = 1, \ldots, n \text{ and}$$

$$T(e_j) = \sum_{i=1}^{m} b_{ij} f_i \ \forall \ j = 1, \ldots, n$$

$\therefore (S+T)(e_j) = S(e_j) + T(e_j)$ (by definition of $S+T$)

$$= \sum_{i=1}^{m} a_{ij} f_i + \sum_{i=1}^{m} b_{ij} f_i$$

$$= \sum_{i=1}^{m} (a_{ij} + b_{ij}) f_i$$

Thus, by definition of the matrix with respect to B_1 and B_2, we get $[S+T] = [a_{ij} + b_{ij}]$.

Now, $(\alpha S)(e_j) = \alpha(S(e_j))$ (by definition of αS)

$$= \alpha\left(\sum_{i=1}^{m} a_{ij} f_i\right)$$

Two matrices can be added if and only if they are of the same size

$$= \sum_{i=1}^{m} (\alpha a_{ij}) f_i$$

Thus, $[\alpha S] = [\alpha a_{ij}]$. ■

We can view an $m \times n$ matrix A as a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$. The matrix of this linear operator with respect to the standard bases in $\mathbb{R}^n$ and $\mathbb{R}^m$ is precisely A. So, Theorem 1 says our definition of matrix addition in Unit 1, Definition 2 is compatible with the addition of linear operator.

Now, let us define the scalar multiple of a matrix, again motivated by Theorem 1.

Definition: Let α be a scalar, i.e., $\alpha \in \mathbf{F}$, and let $A = [a_{ij}]_{m \times n}$. Then we define the **scalar multiple of the matrix $\mathbf{A}$ by the scalar** α to be the matrix

$$\alpha A = \begin{bmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{bmatrix}$$

In other words, αA is the $m \times n$ matrix whose (i, j)th element is α times the (i, j)th element of A.

Again, this agrees with our definition of scalar multiplication in Unit 1.

Remark 1: The way we have defined the sum and scalar multiple of matrices allows us to write Theorem 1 as follows:

$$[\mathbf{S} + \mathbf{T}]_{\mathbf{B}_2}^{\mathbf{B}_1} = [\mathbf{S}]_{\mathbf{B}_2}^{\mathbf{B}_1} + [\mathbf{T}]_{\mathbf{B}_2}^{\mathbf{B}_1}$$

$$[\alpha \mathbf{S}]_{\mathbf{B}_2}^{\mathbf{B}_1} = \alpha [\mathbf{S}]_{\mathbf{B}_2}^{\mathbf{B}_1}.$$

The following exercise will help you in checking if you have understood the contents of Sections 2.2.2 and 2.2.3.

E7) Define $\mathbf{S} : \mathbf{R}^3 \to \mathbf{R}^2 : \mathbf{S}(x, y) = (x, 0, y)$ and $\mathbf{T} : \mathbf{R}^2 \to \mathbf{R}^3 : \mathbf{T}(x, y) = (0, x, y)$. Let B_1 and B_2 be the standard bases for $\mathbf{R}^2$ and $\mathbf{R}^3$, respectively.

Then what are $[S]_{B_2}^{B_1}$, $[T]_{B_2}^{B_1}$, $[S+T]_{B_2}^{B_1}$, $[\alpha S]_{B_2}^{B_1}$, for any $\alpha \in \mathbf{R}$.

We have seen in E28), Unit 2 that $M_{mn}(F)$ is a vector space over F. In the next section we find the dimension of $M_{mn}(F)$ over F.

9.2.2 Dimension of $\mathbf{M_{mn}(F)}$ over $\mathbf{F}$

What is the dimension of $\mathbf{M_{mn}(F)}$ over $\mathbf{F}$? To answer this question we prove the following theorem. But, before you go further, check whether you remember the definition of a vector space isomorphism (Unit 7).

Theorem 2: Let U and V be vector **spaces** over $\mathbf{F}$ of dimensions n and m, respectively. Let B_1 and B_2 be a pair of bases of U and V, respectively. The mapping $\phi : L(U, V) \to \mathbf{M_{mn}(F)}$, given by $\phi(T) = [T]_{B_2}^{B_1}$ is a vector space isomorphism.

Proof: The fact that ϕ is a linear transformation follows from Theorem 1. We proceed to show that the map is also $1-1$ and onto. For the rest of the proof we shall denote $[S]_{B_2}^{B_1}$ by $[S]$ only, and take

$B_1\{e_1, \ldots, e_n\}$, $B_2\{f_1, f_2, \ldots, f_m\}$.

ϕ is $\mathbf{1-1:}$ Suppose, $S, T \in L(U, V)$ be such that $\phi(S) = \phi(T)$.
Then $[S] = [T]$. Therefore, $S(e_j) = T(e_j)\ \forall\ e_j \in B_1$.
Thus, by Unit 5 (Theorem 1), we have $S = T$.

ϕ **is on** $\mathbf{0:}$ If $A \in \mathbf{M_{mn}(F)}$ we want to construct $T \in L(U, V)$ such that $\phi(T) = A$. Suppose $A = [a_{ij}]$. Let $v_1, \ldots, v_n \in V$ such that

$v_j = \sum_{i=1}^{m} a_{ij} f_i$ for $j = 1, \ldots, n$.

Then, by Theorem 3 of Unit 5, there exists as linear transformation

$T \in L(U, V)$ such that $T(e_j) = v_j = \sum_{i=1}^{n} a_{ij} f_i$.

Thus, by definition, $\phi(T) = A$.
Therefore, ϕ is a vector space isomorphism.
A corollary to this theorem gives us the dimension of $\mathbf{M_{mn}(F)}$.

Corollary: Dimension of $\mathbf{M_{mn}(F)} = mn$.

Proof: Theorem 2 tells us that $\mathbf{M_{mn}(F)}$ is isomorphic to $L(U, V)$.
Therefore, $\dim_{\mathbf{F}} \mathbf{M_{mn}(F)} = \dim_{\mathbf{F}} L(U, V)$ (by Theorem 12 of Unit 5) $= mn$, from Unit 6 (Theorem 1).

Why do you think we chose such a roundabout way for obtaining $\dim \mathbf{M_{mn}(F)}$? We could as well have tried to obtain mn linearly independent $m \times n$ matrices and show that they generate $\mathbf{M_{m\times n}(F)}$. But that would be quite tedious (see E16). Also, we have done so much work on $L(U, V)$ so why not use that! And, doesn't the way we have used seem neat?

Now for some exercises related to Theorem 2.

E8) At most, how many matrices can there be in any linearly independent subset of $\mathbf{M_{2\times 3}(F)}$?

E9) Are the matrices $[1, 0]$ and $[1, -1]$ linearly independent over $\mathbf{R}$?

Now we move on to the next section, where we see some ways of getting new matrices from given ones.

9.3 SOME TYPES OF MATRICES

We discussed some special matrices like upper triangular and lower triangular matrices. We now discuss some of these matrices and their relation to linear transformations.

9.3.1 Diagonal Matrix

Let U and V be vector spaces over $\mathbf{F}$ of dimension n. Let $B_1 = \{e_1, \ldots, e_n\}$ and $B_2 = \{f_i, \ldots, f_n\}$ be based of U and V, respectively. Let $d_1, \ldots, d_n \in \mathbf{F}$. Consider the transformation
$T : U \to V : T(a_1e_1 + \cdots + a_ne_n) = a_1d_1f_1 + \cdots + a_nd_nf_n$.
Then $T(e_1) = d_1f_1,\ T(e_2) = d_2f_2,\ \ldots,\ T(e_n) = d_nf_n$.

$$\therefore [T]_{B_2}^{B_1} = \begin{bmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & d_n \end{bmatrix}.$$

Such a matrix is called a diagonal matrix. Let us see what this means.

Let $A = [a_{ij}]$ be a square matrix. The entries $a_{11}, a_{22}, \ldots, a_{nn}$ are called the **diagonal entries** of A. This is because they lie along the diagonal, from left to right, of the matrix. All the other entries of A are called the **off-diagonal entries** of A.

A square matrix whose off-diagonal entries are zero (i.e., $a_{ij} = 0\ \forall\ i \neq j$) is called a **diagonal matrix**. The diagonal matrix

$$\begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & d_n \end{bmatrix}$$

is denoted by $\mathbf{diag\ (d_1, d_2, \ldots, d_n)}$.

Note: The d_i's may or may not be zero. What happens if all the d_i's are zero? Well, we get the $n \times n$ zero matrix, which corresponds to the zero operator.

If $d_i = 1\ \forall\ i = 1, \ldots, n$, we get the **identity matrix**, I_n (or I, when the size is understood).

E10) Show that I, is the matrix associated to the identity operator from $\mathbf{R^n}$ to $\mathbf{R^n}$.

If $\alpha \in \mathbf{F}$, the linear operator $\alpha I : \mathbf{R^n} \to \mathbf{R^n} : \alpha I(v) = \alpha v$, for all $v \in \mathbf{R^n}$, is called a **scalar operator**. Its matrix with respect to any basis is $\alpha I = \text{diag}(\alpha, \alpha, \ldots, \alpha)$. Such a matrix is called a **scalar matrix**. It is a diagonal matrix whose diagonal entries are all equal.

With this much discussion on diagonal matrices, we move onto describe triangular matrices.

9.3.2 Triangular Matrix

Let $B = \{e_1, e_2, \ldots, e_n\}$ be a basis of a vector space V. Let $S \in L(V, V)$ be an operator such that

$$\begin{aligned} S(e_1) &= a_{11}e_1 \\ S(e_2) &= a_{12}e_1 + a_{22}e_2 \\ &\vdots \quad \vdots \quad \vdots \\ S(e_n) &= a_{1n}e_1 + a_{2n}e_2 + \cdots + a_{nn}e_n, \end{aligned}$$

Then, the matrix of S with respect to B is

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Note that $a_{ij} = 0 \ \forall \ i > j$.

A square matrix A such that $a_{ij} = 0 \ \forall \ i > j$ is called an **upper triangular matrix**. If $a_{ij} = 0 \ \forall \ i \geq j$, then A is called **strictly upper triangular**.

For example, $\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ are all upper triangular, while

$\begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}$ is strictly upper triangular.

Note that every strictly upper triangular matrix is an upper triangular matrix.

Now let $T : V \to V$ be an operator such that $T(e_j)$ is a linear combination of $e_j, e_{j1}, \ldots, e_n \forall \ j$. The matrix of T with respect to B is

$$[T]_B = \begin{bmatrix} b_{11} & 0 & 0 & \cdots & 0 \\ b_{21} & b_{22} & \vdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & b_{n3} & \cdots & b_{nn} \end{bmatrix}$$

Note that $b_{ij} = 0 \ \forall \ i < j$.

Such a matrix is called a **lower triangular matrix**. If $b_{ij} = 0$ for all $i \le j$, then B is said to be a **strictly lower triangular matrix**.

The matrix

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 1 & 0 & 5 & 0 \end{bmatrix}$$

is a strictly lower triangular matrix. Of course, it is also lower triangular!

Remark 4: If A is an upper triangular 3×3 matrix, say $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$,

then $A^t = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix}$, a lower triangular matrix.

In fact, for any $n \times n$ upper triangular matrix A, its transpose is lower triangular, and vice-versa.

E11) If an upper triangular matrix A is symmetric, then show that it must be a diagonal matrix.

In the next section, we related the multiplication of matrices to composition of linear transformations.

9.4 MATRIX MULTIPLICATION

9.4.1 Matrix of the Composition of Linear Transformations

Let U, V and W be vector spaces over $\mathbf{F}$, of dimensions p, n and m, respectively. Let B_1, B_2 and B_3 be bases of these respective spaces. Let $T \in L(U, V)$ and $S \in L(V, W)$. Then $ST(= ST) \in L(U, W)$ (see Sec.6.4).

Suppose $[T]_{B_2}^{B_1} = B = [b_{jk}]_{n \times p}$ and $[S]_{B_3}^{B_2} = A = [a_{ij}]_{m \times n}$.

We ask: What is the matrix $[ST]_{B_2}^{B_1}$?

To answer this we suppose

$B_1 = \{e_1, e_2, \ldots, e_p\}$

$B_2 = \{f_1, f_2, \ldots, f_n\}$

$B_3\{g_1, g_2, \ldots, g_m\}$.

Then, we know that $T(e_k) = \sum_{j=1}^{n} b_{jk} f_j \; \forall \, k = 1, 2, \ldots, p$,

and $S(f_j) = \sum_{i=1}^{m} a_{ij} g_i \;\; \forall \;\; j = 1, 2, \ldots, n$.

Therefore,

$$S \circ T(e_k) = S(T(e_k)) = S\left(\sum_{j=1}^{n} b_{jk} f_j\right) = b_{1k}S(f_1) + b_{2k}S(f_2) + \cdots + b_{nk}S(f_n)$$

$$= b_{1k}\left(\sum_{i=1}^{m} a_{ij} g_i\right) + b_{2k}\left(\sum_{i=1}^{m} a_{i2} g_i\right) + \cdots + b_{nk}\left(\sum_{i=1}^{m} a_{in} g_i\right)$$

$$= \sum_{i=1}^{m} (a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{in}b_{nk}) g_i,$$ on collecting the coefficients of g_i.

Thus, $[ST]_{B_3}^{B_1} = [c_{ik}]_{m \times p}$, where $c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk}$

We define the matrix $[c_{jk}]$ to be the product AB.

We now make an observation.

Remark 5: If $T \in L(U, V)$ and $S \in L(V, W)$, then $[ST]_{B_2}^{B_1} = [S]_{B_3}^{B_2}[T]_{B_2}^{B_1}$, where B_1, B_2, B_3 are the bases of $U, V, W,$ respectively.

Let us illustrate this remark.

Example 3: Let $T: \mathbf{R}^2 \to \mathbf{R}^3$ be a linear transformation such that $T(x, y) = (2x + y, x + 2y, x + y)$. Let $S: \mathbf{R}^3 \to \mathbf{R}^2$ be defined by $S(x, y, z) = (-y + 2z, y - z)$. Obtain the matrices $[T]_{B_2}^{B_1}, [S]_{B_1}^{B_2}$, and $[S \circ T]_{B_1}$, and verify that $[S \circ T]_{B_1} = [S]_{B_1}^{B_2}[T]_{B_2}^{B_1}$, where B_1 and B_2 are the standard bases in $\mathbf{R}^2$ and $\mathbf{R}^3$, respectively.

Solution: Let $B_1 = \{e_1, e_2\}, B_2 = \{f_1, f_2, f_3\}$.

Then $T(e_1) = T(1, 0) = (2, 1, 1) = 2f_1 + f_2 + f_3$

$T(e_2) = T(0, 1) = (1, 2, 1) = f_1 + 2f_2 + f_3$

Thus,

$$[T]_{B_2}^{B_1} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}$$

Also,

$S(f_1) = S(1, 0, 0) = (0, 0) = 0.e_1 + 0.e_2$

$S(f_2) = S(0, 1, 0) = (-1, 1) = -e_1 + e_2$

$S(f_3) = S(0, 0, 1) = (2, -1) = 2e_1 + e_2$

Thus,

$$[S]_{B_1}^{B_2} = \begin{bmatrix} 0 & -1 & 2 \\ 0 & 1 & -1 \end{bmatrix}$$

So, $$[S]_{B_1}^{B_2}[T]_{B_2}^{B_1} = \begin{bmatrix} 0 & -1 & 2 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

Also, $S \circ T(x, y) = S(2x + y, x + 2y, x + y)$

$= (-x - 2y + 2x + 2y, x + 2y - x - y)$

$= (x, y).$

Thus, $S \circ T - 1$, the identity map.

This means $[S \circ T]_{B_1} = I_2$.

Hence, $[S \circ T]_{B_1} = [S]_{B_1}^{B_2} [T]_{B_2}^{B_1}$.

E12) Let $S : \mathbf{R^3} \to \mathbf{R^3} : S(x, y, z) = (0, x, y)$, and $T : \mathbf{R^3} \to \mathbf{R^3} : T(x, y, z) = (x, 0, y)$. Show that $[S \circ T]_B = [S]_B [T]_B$, where B is the standard basis of $\mathbf{R^3}$.

E13) Let A, B be two diagonal $n \times n$ matrices over $\mathbf{F}$. Show that AB is also a diagonal matrix.

Now we shall go on to introduce you to the concept of an invertible matrix.

9.5 INVERTIBLE MATRICES

In this section we will relate the invertibility of a linear transformation to the invertibility of its linear transformation with respect to a basis. We will also see how bases of a vector space are related to each other by an invertible matrix.

9.5.1 Inverse of a Matrix

Just as we defined the operations on matrices by considering them on linear operations first, we give a definition of invertibility for matrices based on consideration of invertibility of linear operators.

It may help you to recall what we mean by an invertible linear transformation. A linear transformation $T : U \to V$ is invertible if

a) T is $1-1$ and onto, or, equivalently,

b) there exists a linear transformation $S : V \to U$ such that $S \circ T_U, T \circ SI_V$.

In particular, $T \in L(V, V)$ is said to be invertible if $\exists\, S \in L(V, V)$ such that $ST = TS = I$.

We have the following theorem involving the matrix of an invertible linear operator.

Theorem 3: Let V be an n-dimensional vector space over a field $\mathbf{F}$, and B be a basis of V. Let $T \in L(V,V)$. T is invertible iff there exists $A \in \mathbf{M_n(F)}$ such that $[T]_B A = I_n = A[T]_B$.

Proof: Suppose T is invertible. Then $\exists\, S \in L(V, V)$ such that $TS = ST = I.$ Then, by Theorem 2, $[TS]_B = [ST]_B = I.$ That is, $[T]_B[S]_B = [S]_B[T]_B = I.$ Take $A = [S]_B.$ Then $[T]_B A = I = A[T]_B.$

Conversely, suppose $\exists$ a matrix A such that $[T]_B A = A[T]_B = I.$
Let $S \in L(V, V)$ be such that $[S]_B = A.$ (S exists because of Theorem 2.)
Then $[T]_B[S]_B = [S]_B[T]_B = I = [I]_B.$ Thus, $[TS]_B[ST]_B = [I]_B.$
So, by Theorem 2, $TS = ST = I.$ That is, T is invertible.

We will now make a few observation about the matrix inverse, in the form of a theorem.

Theorem 4: a) If A is invertible, then

i) A^{-1} is invertible and $\mathbf{(A^{-1})^{-1} = A}$.

ii) A^t is invertible and $\mathbf{(A^t)^{-1} = (A^{-1})^t}$.

b) If $A, B \in \mathbf{M_n(F)}$ are invertible, then AB is invertible and $\mathbf{(AB)^{-1} = B^{-1}A^{-1}}$.

Proof: (a) By definition,

$$AA^{-1} = A^{-1}A = I \qquad (1)$$

i) Equation (1) shows that A^{-1} is invertible and $(A^{-1})^{-1} = A.$

ii) If we take transposes in Equation (1) and use the property that $(AB)^t = B^t A^t,$ we get

$$(A^{-1})^t A^t = A^t (A^{-1})^t = I^t = I.$$

So A^t is invertible and $(A^t)^{-1} = (A^{-1})^t.$

(b) To prove this we will use the associativity of matrix multiplication. Now

$$(AB)(B^{-1}A^{-1}) = [A(BB^{-1})]A^{-1} = AA^{-1} = I.$$

$$(B^{-1}A^{-1})(AB) = B^{-1}[(A^{-1}A)B] = B^{-1}B = I.$$

So AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}.$

We now relate matrix invertibility with the linear independence of its rows or columns. When we say that the m rows of $A = [a_{ij}] \in \mathbf{M_{m\times n}(F)}$ are linearly independent, what do we mean? Let $R_1, \ldots, R_m$ be the m row vectors $[a_{11}, a_{12}, \ldots, a_{1n}], [a_{21}, \ldots, a_{2n}], \ldots, [a_{m1}, \ldots, a_{mn}],$ respectively. We say that they are linearly independent if, whenever $\exists\, a_1, \ldots, a_m \in \mathbf{F}$ such that $a_1R_1 + \cdots + a_mR_m = \mathbf{0},$
then $a_1 = 0, \ldots, a_m = 0.$

Similarly, the n columns $C_1, \ldots, C_n$ of A are linearly independent if
$b_1C_1 + \cdots + b_nC_n = 0$
$\Rightarrow b_1 = 0, b_2 = 0, \ldots, b_n = 0,$ where $b_1, \ldots, b_n \in \mathbf{F}.$

We have the following result.

Theorem 5: Let $A \in \mathbf{M_n(F)}$. Then the following conditions are equivalent.

a) A is invertible.
b) The columns of A are linearly independent.
c) The rows of A are linearly independent.

Proof: We first prove (a)$\Leftrightarrow$(b), using Theorem 4. Let V be an n-dimensional vector space over $\mathbf{F}$ and $B = \{e_1, \ldots, e_n\}$ be a basis of V. Let $T \in L(V, V)$ be such that $[T]_B = A$. Then A is invertible iff T is invertible iff $T(e_1), T(e_2), \ldots, T(e_n)$ are linearly independent (see Unit 5, Theorem 9). Now we define the map

$$\theta : V \to M_{n\times 1}(\mathbf{F}) : \theta(a_1e_1 + \cdots + a_ne_n) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Before continuing the proof we give an exercise

E14) Show that θ is a well-defined isomorphism.

Now let us go on with proving Theorem 7.

Let $C_1, C_2, \ldots, C_n$ be the columns of A. Then $\theta(T(e_1)) = C_i$ for all $i = 1, \ldots, n$. Since θ is an isomorphism, $T(e_1), \ldots, T(e_n)$ are linearly independent iff $C_1, C_2, \ldots, C_n$ are linearly independent. Thus, A is invertible iff $C_1, \ldots, C_n$ are linearly independent. Thus, we have proved (a)$\Leftrightarrow$(b).

Now, the equivalence of (a) and (c) follows because A is invertible
$\Leftrightarrow A^t$ is invertible.
$\Leftrightarrow$ the columns of A^t are linearly independent (as we have just shown)
$\Leftrightarrow$ the rows of A are linearly independent (since the columns of A^t are the rows of A).
So we have shown that (a)$\Leftrightarrow$(c).
Thus, the theorem is proved.

From the following example you can see how Theorem 7 can be useful.

Example 4: Let $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \in \mathbf{M_3(R)}$.

Determine whether or not A is invertible.

Solution: Let R_1, R_2, R_3 be the rows of A. We will show that they are linearly independent.

The row operations $R_3 \to R_3 - R_1,\ R_3 \to R_3 - R_2$ gives $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix}$.

Since there are no zero rows, the rows are linearly independent.
Thus, by Theorem 5, A is invertible.

***

E15) Check if $\begin{bmatrix} 2 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 3 & 0 \end{bmatrix} \in \mathbf{M_3(Q)}$ is invertible.

9.5.2 Matrix of Change of Basis

Let V be an n-dimensional vector space over $\mathbf{F}$. Let $B = \{e_1, e_2, \ldots, e_n\}$ and $B'\{e'_1, e'_2, \ldots, e'_n\}$ be two bases of V. Since $e_j \in V$, for every j, it is a linear combination of the elements of B. Suppose,

$$e'_j = \sum_{i=1}^{n} a_{ij} e_i \ \forall \ j = 1, \ldots, n.$$

The $n \times n$ matrix $A = [a_{ij}]$ is called the **matrix of the change of basis from $\mathbf{B}$ to $\mathbf{B}'$**. It is denoted by $\mathbf{M_B^{B'}}$.

Note that A is the matrix of the transformation $T \in L(V, V)$ such that $T(e_j) = e'_j \ \forall \ j = 1, \ldots, n,$ with respect to the basis B. Since $\{e'_1, \ldots, e'_n)$ is a basis of V, from Unit 5 we see that T is $1-1$ and onto. Thus T is invertible. So A is invertible. Thus, **the matrix of the change of basis** from B to B' is invertible.

Note: a) $\mathbf{M_B^B = I_n}$, This is because, in this case $e'_i = e_i \ \forall \ i = 1, 2, \ldots, n.$

b) $M_B^{B'} = [I]_B^{B'}$. This is because

$$I(e'_j) = e'_j = \sum_{i=1} a_{ij} e_i \ \forall \ j = 1, 2, \ldots, n.$$

Now suppose A is any invertible matrix. By Theorem 2, $\exists\ T \in L(V, V)$ such that $[T]_B = A$. Since A is invertible, T is invertible. Thus, T is $1-1$ and onto. Let $f_i = T(e_i) \ \forall \ i = 1, 2, \ldots, n.$ Then $B' = \{f_1, f_2, \ldots, f_n\}$ is also a basis of V, and the matrix of change of basis from B to B' **is** A.

In the above discussion, we have just proved the following theorem.

Theorem 8: Let $B = \{e_1, e_2, \ldots, e_n\}$ be a fixed basis of V. The mapping $B' \to M_B^{B'}$ is a $1-1$ and onto correspondence between the set of all bases of V and the set of invertible $n \times n$ matrices over $\mathbf{F}$.

Let us see an example of how to obtain $M_B^{B'}$.

Example 4: In $\mathbf{R}^2$, $B = \{e_1, e_2\}$ is the standard basis. Let B' be the basis obtained by rotating B through an angle θ in the anti-clockwise direction (see Fig.1). Then $B' = (e_1', e_2')$ where $e_1' = (\cos\theta, \sin\theta), e_2' = (-\sin\theta, \cos\theta)$. Find $M_B^{B'}$.

Fig.1

Solution: $e_1' = \cos\theta(1, 0) + \sin\theta(0, 1)$, and

$$e_2' = -\sin\theta(1, 0) + \cos\theta(0, 1)$$

Thus, $M_B^{B'} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$

***

Try the following exercise.

E16) Let B be the standard basis of $\mathbf{R}^3$ and B' be another basis such that $M_B^{B'} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix}$. What are the elements of B'?

What happens if we change the basis more than once? The following theorem tells us something about the corresponding matrices.

Theorem 9: Let B, B', B'' be three bases of V. Then $\mathbf{M_B^{B'} M_B^{B''}} = \mathbf{M_B^{B'}}$.

Proof: Now, $\mathbf{M_B^{B'} M_B^{B''}} = [I]_B^B, \mathbf{[I]_B^B}$

$$= [I \circ I]_B^B = \mathbf{M_B^B}.$$

An immediate useful consequence is

Corollary: Let B, B' be two bases of V. Then $M_B^{B'} M_{B'}^B = I = M_{B'}^B M_B^{B'}$.
That is, $\left(M_B^{B'}\right)^{-1} = M_{B'}^B$.

Proof: By Theorem 9,

$$\mathbf{M_B^{B'} M_B^B} = \mathbf{M_B^B} = \mathbf{I}$$

Similarly, $\mathbf{M_B^B M_B^B} = \mathbf{M_B^B} = \mathbf{I}$.

But, how does the change of basis affect the matrix associated to a given linear transformation? In Sec.7.2 we remarked that the matrix of a linear transformation depends upon the pair of bases chosen. The relation between the matrices of a transformation with respect to two pairs of bases can be described as follows.

Theorem 10: Let $T \in L(U, V)$. Let $B_1 = \{e_1, \ldots, e_n\}$ and $B_2 = \{f_1, \ldots, f_m\}$ be a pair of bases of U and V, respectively.
Let $B_1' = \{e_1', \ldots, e_n'\}$, $B_2' = \{f_1', \ldots, f_m'\}$ be another pair of bases of U and V, respectively. Then,

$$[T]_{B_2'}^{B_1'} = \mathbf{M}_{B_2'}^{B_2} [\mathbf{T}]_{B_2}^{B_1} \mathbf{M}_{B_1}^{B_1'}.$$

Proof: $[T]_{B_2'}^{B_1'} = [Iv \circ T \circ Iu]_{B_2'}^{B_1'} = [Iv]_{B_2'}^{B_2}[Iu]_{B_1}^{B_1'}$
(where $I_u =$ identity map on U and $Iv =$ identity map on V)

$$= \mathbf{M}_{B_2'}^{B_2} [\mathbf{T}]_{B_2}^{B_1} \mathbf{M}_{B_1}^{B_1'}$$

Now, a corollary to Theorem 10, which will come in handy in the next block.

Corollary: Let $T \in L(V, V)$ and B, B' be two bases of V. Then $[T]_B = P^{-1}[T]_B P$, where $P = \mathbf{M}_B^{B'}$.

Proof: $[T]_B = \mathbf{M}_B^{B'}[T]_B \mathbf{M}_B^{B'} = P^{-1}[T]_B P$, by the corollary to Theorem 9.

We can compute the matrix $M_B^{B'}$ using row reduction when B and B' are subsets of $\mathbb{R}^n$. Let us look at an example.

Example 5: Let $B = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$ and $B' = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$. Find $M_B^{B'}$.

Solution: We have

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = a_{11} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + a_{21} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + a_{31} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \end{bmatrix} = \overline{B} \begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix}.$$

Here we write $\overline{B}$ for the matrix whose columns are the elements of B. Since B is a linearly independent set, the columns of $\overline{B}$ are linearly independent. So, $\overline{B}$ is invertible. Therefore

$$\begin{bmatrix} a_{11} \\ a_{21} \\ a_{31} \end{bmatrix} = \overline{B}^{-1} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Similarly, we can write down the equations

$$\begin{bmatrix} a_{12} \\ a_{22} \\ a_{32} \end{bmatrix} = \overline{B}^{-1} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} a_{13} \\ a_{23} \\ a_{33} \end{bmatrix} = \overline{B}^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}.$$

$$\therefore M_B^{B'} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \overline{B}^{-1}\overline{B}'$$

where $\overline{B}'$ is the matrix formed by the rows of B'. To find $\overline{B}^{-1}B'$ we set up a 3×6 matrix $C = [\overline{B} \mid \overline{B}']$. We then use row operations to reduce the matrix formed by the first three column to identity. We get $[I \mid D]$ when $D = \overline{B}^{-1}\overline{B}'$. So, we form the matrix

$$C = \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \end{array}\right].$$

Carrying out row operations $R_2 \to R_2 - R_1$, $R_3 \to R_3 - R_1$, $R_2 \leftrightarrow R_3$, $R_2 \to -R_2, R_1 \to R_1 - R_2$, $R_3 \to -R_3$, $R_2 - R_2 - R_3$, we get

$$C' = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & -1 \end{array}\right].$$ (Notice that C' is the RREF of C.)

Therefore $M_B^{B'} \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 1 \\ 1 & 0 & -1 \end{bmatrix}$.

***

In general, if $B = \left\{ \begin{bmatrix} b_{11} \\ b_{21} \\ \vdots \\ b_{n1} \end{bmatrix}, \begin{bmatrix} b_{12} \\ b_{22} \\ \vdots \\ b_{n2} \end{bmatrix}, \ldots, \begin{bmatrix} b_{1n} \\ b_{2n} \\ \vdots \\ b_{nn} \end{bmatrix} \right\}$

and

$$B' \left\{ \begin{bmatrix} b'_{11} \\ b'_{21} \\ \vdots \\ b'_{n1} \end{bmatrix}, \begin{bmatrix} b'_{12} \\ b'_{22} \\ \vdots \\ b'_{n2} \end{bmatrix}, \ldots, \begin{bmatrix} b'_{1n} \\ b'_{2n} \\ \vdots \\ b'_{nn} \end{bmatrix} \right\},$$

then, as in Example ? we have $M_B^{B'} = \overline{B}^{-1}\overline{B}'$. To compute $M_B^{B'}$, we set up the matrix $[\overline{B} \mid \overline{B}']$ and reduce this to RREF $C' = [I_n \mid D]$.
Then, $M_B^{B'} = D$.

Try the exercises below to check your understanding of the example?

E17) For the following pairs of basis B and B', find the change of base matrix $M_B^{B'}$.

i) $B = \{(1, -1), (2, 1)\}$, $B'\{(1, 1), (1, 2)\}$.

ii) $B'\{(1, -1, 1), (-1, 0, 1), (1, 1, 1)\}$, $B' = \{(1, -1, 1), (1, 1, -1), (0, 0, 1)\}$.

Example 6: Let $B = \{(1,-1,-1),(1,1,0),(1,0,1)\}$ and $B' = \{(1,1,-1),(0,1,1),(0,0,1)\}$. Let $T:\mathbb{R}^3 \to \mathbb{R}^3$ be a linear operator such that $[T]_B = \begin{bmatrix} 1 & 1 & 2 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. Find $[T]_{B'}$.

Solution: By Corollary to Theorem 10, we have $[T]_{B'} = P^{-1}[T]_B P$ where $P = M_B^{B'}$. We compute $M_B^{B'}$ as before. We set

$$C = \left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 1 & 0 \\ -1 & 0 & 1 & -1 & 1 & 1 \end{array}\right].$$ Carrying out row operators

$R_2 \to R_2 + R_1,\ R_3 \to R_3 + R_1,\ R_2 \to \dfrac{R_2}{2}, R_1 \to R_1 - R_2, R_3 \to R_3 - R_2,$ we get the RREF of C is $C' = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -1/2 & 0 \\ 0 & 1 & 0 & 1 & 1/2 & 0 \\ 0 & 0 & 1 & -1 & 1/2 & 1 \end{array}\right]$.

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

We row find P^{-1} using row reduction. We set

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

Using row operations $R_1 \leftrightarrow R_2,\ R_3 \to R_3 + R_1,$

$R_2 = (-2)R_2,\ R_1 \to R_1 - \dfrac{R_2}{2},\ R_3 \to R_3 - R_2,$

We get $\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & -2 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 1 \end{bmatrix}$. $\therefore P^{-1} = \begin{bmatrix} 1 & 1 & 0 \\ -2 & 0 & 0 \\ 2 & 1 & 1 \end{bmatrix}$.

$$\therefore [T]_{B'} = P^{-1}[T]_B P = \begin{bmatrix} 1 & 1 & 0 \\ -2 & 0 & 0 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 2 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1/2 & 0 \\ 1 & 1/2 & 0 \\ -1 & 1/2 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 2 & 2 \\ 2 & -4 & -4 \\ -1 & 5 & 5 \end{bmatrix}.$$

***

Here are some exercises for you to try.

E18) Let $T:\mathbb{R}^2 \to \mathbb{R}^2$ be such that its matrix with respect to the standard basis is $\begin{bmatrix} 1 & 6 \\ 2 & 2 \end{bmatrix}$. Find the matrix with respect to the basis $\{(-2,1),(3,2)\}$.

E19) Let $T:\mathbb{R}^3 \to \mathbb{R}^3$ be such that its matrix with respect to the basis $B = \{(1, 0, 1), (-1, 0, 1), (0, 1, 1)\}$ is $\begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$. Find the matrix of the linear transformation with respect to the basis $B' = \{(1, -1, 1), (1,1,1), (0, 0, 1)\}$.

Let us now recapitulate all that we have covered in this unit.

2.7 SUMMARY

We briefly sum up what has been done in this unit.

1. We defined matrices and explained the method of associating matrices with linear transformations.

2. We showed what we mean by sums of matrices and multiplication of matrices by scalars.

3. We proved that $\mathbf{M_{m\times n}(F)}$ is a vector space of dimension mn over $\mathbf{F}$.

4. We defined the transpose of a matrix, the conjugate of a complex matrix, the conjugate transpose of a complex matrix, a diagonal matrix, identity matrix, scalar matrix and lower and upper triangular matrices.

5. We defined the multiplication of matrices and showed its connection with the composition of linear transformations. Some properties of the matrix product were also listed and used.

6. The concept of an invertible matrix was explained.

7. We defined the matrix of a change of basis, and discussed the effect of change of bases on the matrix of a linear transformation.

2.8 SOLUTIONS/ANSWERS

E1) Suppose $B_1' = \{(1, 0, 1), (0, 2, -1), (1, 0, 0)\}$ and $B_2' = \{(0, 1), (1, 0)\}$

Then $T(1, 0, 1) = (1, 0) = 0.(0, 1) + .(1, 0)$

$T(0, 2, -1) = (0, 2) = 2.(0, 1) + 0, (1, 0)$

$T(1, 0, 0) = (1, 0) = 0.(0, 1) + 1.(1, 0).$

$$\therefore [T]_{B_2'}^{B_1'} = \begin{bmatrix} 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

E2) $B_1 = \{e_1, e_2, e_3\}$, $B_2 = \{f_1, f_2\}$ are the standard bases (given in Example 3).

$T(e_1) = T(1, 0, 0) = (1, 2) = f_1 + 2f_2$

$T(e_2) = T(0, 1, 0) = (2, 3) = 2f_1 + 3f_2$

$T(e_3) = T(0, 0, 1) = (2, 4) = 2f_1 + 4f_2$

$\therefore [T]_{B_2}^{B_1} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 3 & 4 \end{bmatrix}.$

E3) $T(1, 0, 0) = (1, 2) = 1.(1, 2) + 0.(2, 3)$

$T(0, 1, 0) = (2, 3) = 0.(1, 2) + 1.(2, 3)$

$T(1, -2, 1) = (-1, 0) = 3(1, 2) - 2(2, 3)$

$\therefore [T]_{B_2'}^{B_1'} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \end{bmatrix}.$

E4) Let $B = \{1, t, t^2, t^3\}$. Then

$D(1) = 0 = 0.1 + 0.t + 0.t^2 + 0.t^3$

$D(t) = 1 = 1.1 + 0.t + 0.t^2 + 0.t^3$

$D(t^2) = 2t = 0.1 + 2.t + 0.t^2 + 0.t^3$

$D(t^3) = 3t^2 = 0.1 + 0.t + 3.t^2 + 0.t^3.$

Therefore $[D]_B$ is the given matrix.

E5) We now that

$T(e_1) = f_1$

$T(e_2) = f_1 + f_2$

$T(e_3) = f_2$

Therefore, for any $(x, y, z) \in \mathbf{R}^3$

$$\begin{aligned} T(x, y, z) &= T(xe_1 + ye_2 + ze_3) = xT(e_1) + yT(e_2) + zT(e_3) \\ &= xf_1 + y(f_1 + f_2) + zf_2 = (x + y)f_1 + (y + z)f_2 \\ &= (x + y, y + z) \end{aligned}$$

That is, $T : \mathbf{R}^3 \to \mathbf{R}^2 : T(x, y, z) = (x + y, y + z).$

E6) We are given that

$T(1) = 0.1 + 1.i = i$

$T(i) = (-1).1 + 0.1 = -1$

$\therefore$, for any $a + ib \in \mathbf{C}$, we have

$T(a + ib) = aT(1) + bT(i) = ai - b$

E7) $B_1 = \{(1, 0), (0, 1)\}$, $B_2 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$

Now $S(1, 0) = (1, 0, 0)$

$S(0, 1) = (0, 0, 1)$

$\therefore [S]_{B_1 . B_2} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$, a 3×2 matrix

Again, $T(1, 0) = (0, 1, 0)$

$T(0, 1) = (0, 0, 1)$

$\therefore [T]_{B_1.B_2} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$, a 3×2 matrix

$\therefore [S+T]_{B_1.B_2} = [S]_{B_1.B_2} + [T]_{B_1.B_2} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 2 \end{bmatrix}$, and

$[\alpha S]_{B_1.B_2} = \alpha [S]_{B_1.B_2} = \alpha \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & 0 \\ 0 & \alpha \end{bmatrix}$, for any $\alpha \in \mathbf{R}$.

E8) Since $\dim \mathbf{M}_{2\times 3}(\mathbf{R})$ is 6, any linearly independent subset can have 6 elements, at most.

E9) Let $\alpha, \beta \in \mathbb{R}$ such that $\alpha[1, 0] + \beta[1, -1] = [0, 0]$.
Then $[\alpha + \beta, -\beta] = [0, 0]$. Thus, $\beta = 0, \alpha = 0$.
$\therefore$ the matrices are linearly independent.

E10) $I : \mathbf{R^n} \to \mathbf{R^n} : I(x_1, \ldots, x_n) = (x_1, \ldots, x_n)$.
Then, for any basis $B = \{e_1, \ldots, e_n\}$ for $\mathbf{R^n}$, $I(e_i) = e_i$.

$$\therefore [I]_B = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I_n$$

E11) Since A is upper triangular, all its elements below the diagonal are zero. Again, since $A = A^t$, a lower triangular matrix, all the entries of A above the diagonal are zero. $\therefore$, all the off-diagonal entries of A are zero. $\therefore A$ is a diagonal matrix.

E12) $[S]_B = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $[T]_B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

$$\therefore [S]_B [T]_B = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Also, $[S \circ T] = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = [S]_B [T]_B$

E13) Let $A = \text{diag}\,(d_1, \ldots, d_n)$, $B = \text{diag}\,(e_1, \ldots, e_n)$. Then

$$AB = \begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \vdots & \cdots & d_n \end{bmatrix} \begin{bmatrix} e_1 & 0 & 0 & \cdots & 0 \\ 0 & e_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \vdots & \cdots & e_n \end{bmatrix}$$

$$= \begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2e_2 & 0 & \cdots & 0 \\ 0 & 0 & d_3e_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & d_ne_n \end{bmatrix}$$

$$= \text{diag}\,(d_1e_1, d_2e_2, \ldots, d_ne_n).$$

E14) Firstly, θ is a well defined map. Secondly, check that $\theta(v_1 + v_2) = \theta(v_1) + \theta(v_2)$, and $\theta(\alpha v) = \alpha\theta(v)$ for $v, v_1, v_2 \in V$ and $\alpha \in \mathbf{F}$. Thirdly, show that $\theta(v) = 0 \Rightarrow 0$, that is θ is $1-1$. Then, by Unit 5 (Theorem 10), you have shown that θ is an isomorphism.

E15) We will show that its rows are linearly independent over $\mathbb{Q}$.

Carrying out row operations $R_1 \to \frac{R_1}{2}, R_2 \leftrightarrow R_3, R_2 \to \frac{R_2}{3}$, we get

$$\begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Since there are no zeros, the rows are linearly independent.
$\therefore$ the rows of the given matrix is linearly independent.

E16) Let $B = \{e_1, e_2, e_3\}, B' = \{f_1, f_2, f_3\}$. Then

$$f_1 = 0e_1 + 1e_2 + 0e_3 = e_2$$
$$f_2 = e_1 + e_2$$
$$f_3 = e_1 + 3e_3$$
$$\therefore B' = \{e_2, e_1 + e_2, e_1 + 3e_3\}.$$

E17) i) Looking at the elements of B and B' as column vectors, we have $\overline{B} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}, \overline{B'} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$. Setting up $C = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 1 \\ -1 & 1 & 1 & 2 \end{array}\right]$ and finding its RREF, we get $C' = \left[\begin{array}{cc|cc} 1 & 0 & -1/3 & -1 \\ 0 & 1 & 2/3 & 1 \end{array}\right]$.

Therefore $M_B^{B'} = \begin{bmatrix} -1/3 & -1 \\ 2/3 & 1 \end{bmatrix}$.

ii) Looking at the elements of B and B' as column vectors we have $\overline{B} = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \overline{B'} = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 1 \end{bmatrix}$.

Setting up $C = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array}\right]$ and finding its RREF,

we get $C' = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/2 & -1/2 & 1/4 \\ 0 & 1 & 0 & -1 & -1 & 1/2 \\ 0 & 0 & 1 & -1/2 & 1/2 & 1/4 \end{array}\right]$. The row operations are

$R_2 \to R_2 + R_1,\ R_3 \to R_3 - R_1,\ R_2 = -R_2,\ R_1 \to R_1 + R_2,$

$R_3 \to R_3 - 2R_2,\ R_3 \to \dfrac{R_3}{4},\ R_1 \to R_1 + 2R_3.$

Therefore, $M_B^{B'} = \begin{bmatrix} 1/2 & -1/2 & 1/4 \\ -1 & -1 & 1/2 \\ -1/2 & 1/2 & 1/4 \end{bmatrix}$.

E18) We have $B = \{(1, 0), (0, 1)\},\ B' = \{(-2, 1), (3, 2)\}$.

$P = \overline{B}^{-1}\overline{B} = \begin{bmatrix} -2 & 3 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \dfrac{1}{7}\begin{bmatrix} -2 & 3 \\ 1 & 2 \end{bmatrix}.$

$P^{-1} = \left(\dfrac{1}{7}\begin{bmatrix} -2 & 3 \\ 1 & 2 \end{bmatrix}\right)^{-1} = \begin{bmatrix} -2 & 3 \\ 1 & 2 \end{bmatrix}.$

$$P = \begin{bmatrix} 1 & 6 \\ 2 & 2 \end{bmatrix} P^{-1} = \frac{1}{7}\begin{bmatrix} -2 & 3 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 6 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} -2 & 3 \\ 1 & 2 \end{bmatrix} = \frac{1}{7}\begin{bmatrix} -2 & 3 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 4 & 15 \\ -2 & 10 \end{bmatrix}$$

$$= \frac{1}{7}\begin{bmatrix} -14 & 0 \\ 0 & 35 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 0 \\ 0 & 5 \end{bmatrix}.$$

E19) We have $[T]_{B'} = P^{-1}[T]_B P$ where $P = \overline{B}^{-1}\overline{B}$.

We form the matrix $C = \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 1 & -1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right]$. Row reducing using

row operations $R_2 \to R_2 + R_1,\ R_3 \to R_3 - R_1,\ R_2 \to \dfrac{R_2}{2},\ R_1 \to R_1 - R_2$

gives

$$C' = \begin{bmatrix} 1 & 0 & 0 & 1/2 & -1/2 & -1/2 \\ 0 & 1 & 0 & 1/2 & -1/2 & 1/2 \\ 0 & 0 & 1 & 0 & 2 & 1 \end{bmatrix}$$

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

To find P^{-1}, we set up the matrix

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

Row reducing using row operations $R_1 \to 2R_1,\ R_2 \to \frac{R_1}{2},\ R_2 \leftrightarrow R_3,$

$R_2 \to \frac{R_2}{2},\ R \to R_1 + R_2,\ R_1 \to R_1 + \frac{R_3}{2},\ R_2 \to R_2 - \frac{R_3}{2}$ gives

$$\begin{bmatrix} 1 & 0 & 0 & 3/2 & 1/2 & 1/2 \\ 0 & 1 & 0 & 1/2 & -1/2 & 1/2 \\ 0 & 0 & 1 & -1 & 1 & 0 \end{bmatrix}.\quad P^{-1} = \begin{bmatrix} 3/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \\ -1 & 0 & 0 \end{bmatrix}.$$

$$\therefore [T]_{B'} = P^{-1}[T]_B P = \begin{bmatrix} 1 & 6 & 2 \\ -1/2 & 7/2 & 3/2 \\ 3/2 & -11/2 & -3/2 \end{bmatrix}.$$

MISCELLANEOUS EXAMPLES AND EXERCISES

Example 1: In this example, we revisit the projection operator and the reflection operator. Let V be a vector space and W_1 and W_2 be subspaces of V such that $V = W_1 \oplus W_2$. Then, we define the projection of W_1 along W_2 as follows: Let $v \in V$. Then, $v = w_1 + W_2$, $w_1 \in W_1$ and $w_2 \in W_2$. Here w_1 and w_2 are uniquely determined by $v \in V$: we define the projection P_{W_1, W_2} by $P_{W_1, W_2}(v) = w_1$. We define the reflection of W_1 along W_2 by $R_{W_1, W_2}(v) = W_1 - w_2$. By the uniqueness of decomposition of v as $w_1 + w_2$, these functions are well defined. Once we fix the subspaces W_1 and W_2, we will drop the subscripts and simply write P and R for the projection and reflection operators.

1) Check that P and R are linear operators.
2) Check that $P^2 = P$, $R^2 = I$, the identity operator.

Solution: 1) Let $v, v' \in V$. Suppose $v = w_1 + w_2$ and $v' = w_1' + w_2'$.
Then, $P(v) = w_1$, $P(v') = w_1'$. We have $V + v' = (w_1 + w_1') + (w_2 + w_2')$ where $w_1 + w_1' \in W$, and $w_2 + w_2 \in W_2$ since W_1 and W_2 are subspaces of V. Therefore, $P(v + v') = W_1 + w_1' = P(v) + P(v')$. Let $\alpha \in F$, $v \in V$, $v = W_1 + w_2$. Then, $P(V) = W_1$. We have $\alpha v = \alpha w_1 + \alpha w_2$ where $\alpha w_1 \in W_1$ and $\alpha w_2 \in W_2$ since W_1 and W_2 are subspaces. Therefore $P(\alpha v) = \alpha W_1 = \alpha P(v)$. Therefore P is a linear operator.

We have $R(v) = w_1 - w_2$, $R(v') = w_1' - w_2'$.

$$R(v + v') = R((w_1 + w_1') + (w_2 + w_2')) = ((w_1 + w_1') - (w_2 + w_2'))$$
$$= (w_1 - w_1') + (w_2 - w_2') = R(v) + R(v').$$

If $\alpha \in F$, $\alpha v = (\alpha w_1 + \alpha w_2)$.

$\therefore R(\alpha v) = \alpha w_1 - \alpha w_2 = \alpha(w_1 - w_2) = \alpha R(v)$.

Therefore R is a linear operator.

2) We have $P(v) = w_1 = w_1 + 0$. $P(P(v)) = P(w_1 + 0) = w_1$. Therefore $P^2 = P$.
Also, $R(v) = w_1 - w_2 = w_1 + (-w_2)$.
$R^2(v) = R^2(w_1 + (-w_2)) = w_1 - (-w_2) = w_1 + w_2 = v$. In other words $R^2(v) = v$ for all $v \in V$ and therefore $R^2 = I$.

Try the next exercise to check your understanding of the operators R and P.

E1) Shwo that $R = 2, P - I$.

Let us suppoe that $V = \mathbb{R}^2$ or $\mathbb{R}^3$ so that we have a dot product defined on V. Then, we call the decomposition $V = W_1 \oplus W_2$ an **orthogonal** decomposition if $w_1 \cdot w_2 = 0$ for all $w_1 \in W, w_2 \in W_2$. In this situation we call P the orthogonal projection on W_1.

For example, if we take $W_1 = [\{(1, 0)\}], W_2 = [\{(0, 1)\}]$, any $w_1 = W_1$ is of the form $(x, 0)$ and any element $w_2 \in W_2$ is of the for $(0, x)$. Therefore $w_1 \cdot w_2 = 0$. We can write $(x, y) \in \mathbb{R}^2$ as $(x, 0) + (0, y)$. So, $P((x, y)) = (x, 0)$ and $R((x, y)) = (x, 0) - (0, y) = (x, -y)$. Thus there are the usual projection on the x-axis and reflection about x-axis.

Here is an exercise for you.

E2) Describe the projection on y-axis and reflection about y-axis as orthogonal projection and reflection operators with respect to a suitable choice of W_1 and W_2.

E3) Let $m \neq 0$ and $W_1 = \{(x, y) \in \mathbb{R}^2 \mid y = mx\}$ and $W_2 = \{(x, y) \in \mathbb{R}^2 \mid y = -\frac{1}{m}x\}$.

i) Check that $W_1[\{1, m)\}], W_2[\{-m, 1)\}]$.

ii) Prove that $\left\{\frac{(1, m)}{\sqrt{m^2+1}}, \frac{(-m, 1)}{\sqrt{m^2+1}}\right\}$ is an orthonormal basis for $\mathbb{R}^2$.

iii) Prove that $\mathbb{R}^2 = W_1 \oplus W_2$ is an orthogonal decomposition of $\mathbb{R}^2$.

iv) Compute projection and reflection operators with respect to this decomposition.

Example 2: Let V be a vector space over a field F. Let $T: V \to V$ be a linear operator and $W \subseteq V$ be a subspace of V. We call W a T-invariant subspace if $Tw \in W$ for all $w \in W$.

i) Show that $\{0\}, V, \text{Ker}(T)$ and $R(T)$ are all T-invariant subspaces.

ii) If $T: V \to V$ is a linear operator, then every subspace of T is a T-invariant subspace iff there is a $\lambda \in F$ such that $T(v) = \lambda v$ for all $v \in V$.

Solution: i) $T(0) = 0$ so, $\{0\}$ is T-invariant.

If $v \in \text{Ker}(T)$. We have $T(v) = 0$. So, $T(T(v)) = T(0) = 0$ or $T(v) \in \text{Ker}(T)$. Therefore $\text{Ker}(T)$ is T-invariant.

If $v \in R(T), Tv \in R(T)$ by the definition of $R(T)$. So, $R(T)$ is T-invariant

ii) Suppose that there is a $\lambda \in F$ such that $T(v) = \lambda v$ for all $v \in V$. If $v \in W, \lambda v \in W$ since W is a subspace of V.

Let us now prove the converse. Let $v \neq 0$ be in V. Then, $W[\{v\}]$ is a subspace of V. Then, $T(v) \in W$, therefore $T(v) = \lambda_0 v$. We claim that $T(v) = \lambda_0 v$ for all $v \in V$. Let $v' \in V, v' \neq v$. If $v' = \alpha v, T(v') = T(\alpha v) = \alpha T(v) = \alpha\lambda_0 v = \lambda_0(\alpha v) = \lambda_0 v'$.

So, suppose that there is no α such that $v' = \alpha v$.

Taking $W = [\{v'\}]$, there is $\lambda_1 \in F$ such that $T(v') = \lambda_1 v$.

Consider $u = v + v'$. Then $u \neq 0$; otherwise $v' = (-1)v, v' = \alpha v$ with $\alpha = -1$.

There is a λ_3 such that $T(u) = \lambda_3 u$.

$\therefore \lambda_3(v + v') = T(u) = T(v + v') = \lambda_0 v + \lambda_1 v'$.

Since there is no α such that $v' = \alpha v$, v and v' are linearly independent. From $\lambda_3 v + \lambda_3 v' - \lambda_0 v + \lambda_1 v'$, we get $\lambda_1 = \lambda_0 = \lambda_3$ and we are done.

SOLUTIONS/ANSWERS

E1) We have
$2P(v) - I(v) = 2P(w_1 + w_2) - (w_1 + w_2) = 2w_1 - w_1 - w_2 = w_1 - w_2$.

E2) We take $W_1\{, y) \mid y \in \mathbb{R}\}$ and $W_2 = \{(x, 0) \mid x \in \mathbb{R}\}$. Then $P((x, y)) = P((0, y) + (x, 0)) = (0, y)$ and $R((x, y)) = (, y) - (x, 0) = (-x, 0)$. These are, respectively the projection on the y-axis and reflection about the y-axis.

E3) 1) We have $W_1 = \{(x, mx) \mid x \in \mathbb{R}\}$. So, $W_1 = [\{(1, m)\}]$, $W_2 = \{(-mx, x) \mid x \in \mathbb{R}\}$ therefore $W_2 = [\{(-m, 1)\}]$.

2) Computing the dot product, we get
$$\frac{(1, m)}{\sqrt{m^2+1}} \cdot \frac{(-m, 1)}{\sqrt{m^2+1}} = \frac{-m+m}{m^2+1} = 0.$$
So, $\left\{\frac{(1, m)}{\sqrt{m^2+1}}, \frac{(-m, 1)}{\sqrt{m^2+1}}\right\}$ is an orthonormal basis.

3) $\frac{(1, m)}{\sqrt{m^2+1}} \in W_1$ and $\frac{(-m, 1)}{\sqrt{m^2+1}} \in W_2$.

Further, if $\alpha\frac{(1, m)}{\sqrt{m^2+1}} + \beta\frac{(-m, 1)}{\sqrt{m^2+1}} = 0$. Taking dot products sides with $\frac{(1, m)}{\sqrt{m^2+1}}$, we get
$$\alpha\left(\frac{1 \cdot 1 + m \cdot m}{m^2+1}\right) + \beta\left(\frac{1 \cdot (-m) + 1 \cdot m}{(m^2+1)}\right) = 0 \text{ or } \alpha = 0.$$

Similarly taking dot products both sides with $\frac{(-m, 1)}{m^2+1}$, we get $\beta = 0.$

So, $\left\{\frac{(1, m)}{\sqrt{m^2+1}}, \frac{(-m, 1)}{\sqrt{m^2+1}}\right\}$ is a linearly independent set in $\mathbb{R}^2$ and therefore forms a basis for $\mathbb{R}^2$. So, we can write any $v \in \mathbb{R}^2$ as $v = \alpha\left(\frac{(1, m)}{\sqrt{m^2+1}}\right) + \beta\left(\frac{(-m, 1)}{\sqrt{m^2+1}}\right)$. Since $\frac{(1, m)}{\sqrt{m^2+1}} \in W_1, \frac{(-m, 1)}{\sqrt{m^2+1}} \in W_2, \mathbb{R}^2 = W_1 + W_2.$

If $v \in W_1 \cap W_2, v = \alpha\left(\frac{(1, m)}{\sqrt{m^2+1}}\right) = \beta\left(\frac{(-m, 1)}{\sqrt{m^2+1}}\right).$

Taking dot product with $\frac{(1, m)}{\sqrt{m^2+1}}$ both sides, we get

$\alpha\left(\frac{1\cdot 1 + m\cdot m}{m^2+1}\right) = \beta\left(\frac{1\cdot(-m) + m\cdot 1}{m^2+1}\right) = 0$ or $\alpha = 0.$ $\therefore v = 0.$

$\mathbb{R}^2 = W_1 \oplus W_2.$

If $w_1 \in W_1$ and $w_2 \in W_2$, we get $w_1 = \frac{\alpha(1, m)}{\sqrt{m^2+1}}, w_2 = \frac{\beta(-m, 1)}{\sqrt{m^2+1}}.$

$\therefore w_1 \cdot w_2 = 0$ and $\mathbb{R}^2 = W_1 \oplus W_2$ is an orthogonal decomposition.

4) If $(x, y) \in \mathbb{R}^2$, we know that $(x, y) = (u\cdot(x, y))u + (v\cdot(x, y))v$ where $u = \frac{(1, m)}{\sqrt{m^2+1}}, v = \frac{(-m, 1)}{\sqrt{m^2+1}}$. (see Section 2.5).

$u\cdot(x, y) = \frac{x+my}{\sqrt{m^2+1}}, v\cdot(x, y) = \frac{y-mx}{\sqrt{m^2+1}}.$

$\therefore (x, y) = \left(\frac{x+my}{\sqrt{m^2+1}}\right)u + \left(\frac{y-mx}{\sqrt{m^2+1}}\right)v$ is of the form $w_1 + w_2$ with $w_1 \in W_1$ and $w_2 \in W_2$ since $\frac{x+my}{\sqrt{m^2+1}} u \in W_1$ and $\frac{y-m}{\sqrt{m^2 1}} v \in W_2.$

$$\therefore P((x, y)) = \frac{x+my}{\sqrt{m2+1}} u = \frac{x+my(1, m)}{m^2+1} = \left(\frac{x+my}{m^2+1}, \frac{mx+m^2y}{m^2+1}\right)$$

$$R((x, y)) = w_1 - w_2 = \frac{x+my}{\sqrt{m^2+1}}\frac{(1, m)}{\sqrt{m^2+1}} - \frac{y-mx}{\sqrt{m^2+1}}\frac{(-m, 1)}{\sqrt{m^2+1}}.$$

$$= \left(\frac{x+my}{(m^2+1)}, \frac{mx+m^2y}{m^2+1}\right) - \left(\frac{m^2x-my}{m^2+1}, \frac{y-mx}{m^2+1}\right)$$

$$= \left(\frac{(1-m^2)x+2my}{m^2+1}, \frac{(m^2-1)y+2mx}{m^2+1}\right).$$