Chapter 9

Matrix Representation of a Linear Transformation

We restrict attention to finitely generated vector spaces over a fixed field $\mathbb{F}$ and exploit the fact that every finitely generated vector space admits a basis to introduce computational techniques.

Let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be a basis for the vector space V and $\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ for W. Let $T\colon V \to W$ be a linear transformation.

Each $\mathbf{x} \in V$ can be expressed as a linear combination of $\mathbf{e}_1, \dots, \mathbf{e}_n$ in precisely one way, say as,

$$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n = \sum_{j=1}^{n} x_j\mathbf{e}_j \tag{9.1}$$

with $x_j \in \mathbb{F} \quad (j = 1, \dots, n)$.

Similarly, each $\mathbf{y} \in W$ can be written uniquely as

$$\mathbf{y} = y_1\mathbf{f}_1 + \cdots + y_m\mathbf{f}_m = \sum_{i=m}^{n} y_i\mathbf{f}_i \tag{9.2}$$

with $y_i \in \mathbb{F} \quad (i = 1, \dots, m)$.

Put $\mathbf{y} = T(\mathbf{x})$. Then, since T is a linear transformation,

$$\mathbf{y} = T(\mathbf{x}) = T(\sum_{j=1}^{n} x_j\mathbf{e}_j) = \sum_{j=1}^{n} x_jT(\mathbf{e}_j). \tag{9.3}$$

This means T is completely determined by $T(\mathbf{e}_j) \quad (j = 1, \dots, n)$.

Since $T(\mathbf{e}_j) \in W \quad (j = 1, \dots, n)$,

$$T(\mathbf{e}_j) := \sum_{i=1}^{m} a_{ij}\mathbf{f}_i \tag{9.4}$$

for suitable (uniquely determined) $a_{ij} \in \mathbb{F} \quad (i = 1, \dots, m, \ j = 1, \dots, n)$.

It now follows from Equation (9.3) that

$$\begin{aligned}\sum_{i=m}^{n} y_i\mathbf{f}_i &= \sum_{j=1}^{n} x_j(\sum_{i=1}^{m} a_{ij}\mathbf{f}_i) \\ &= \sum_{i=1}^{m}(\sum_{j=1}^{n} x_ja_{ij})\mathbf{f}_i\end{aligned}$$

By uniqueness in Equation (9.2), we deduce that

$$y_i = \sum_{j=1}^{n} x_j a_{ij} = \sum_{j=1}^{n} a_{ij} x_j. \tag{9.5}$$

We use matrix notation to rewrite this.

Since we have fixed a basis for V and a basis for W, we can use the uniqueness of the expressions in Equations (9.1), (9.2) and (9.4) to represent $\mathbf{x} \in V, \mathbf{y} \in W$ and $T : V \to W$ by

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \qquad \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

respectively. This allows us to rewrite Equation 9.5 in terms of matrices as

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{9.6}$$

Observation 9.1. For the moment, this is just notation and nothing else. Think of it as merely a way of storing the data from which the linear transformation can be reconstructed. No algebraic operation has been defined here, even if the reader correctly anticipates further developments. It is important to realise that, at this stage, this representation as nothing more than a convenient notational convention.

Definition 9.2.

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

is the *coordinate vector of the vector* $\mathbf{x} \in V$ *with respect to the basis* $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ *for* V and

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

is the *the matrix of the linear transformation* $T : V \longrightarrow W$ *with respect to the bases* $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ *for* V *and* $\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ *for* W.

Observation 9.3. When $W = V$, we assume that the same basis has been chosen in the codomain as in the domain, unless otherwise specified.

Theorem 9.4. *The matrix of* $id_V : V \to V$ *with respect to any basis for* V *is*

$$\underline{\mathbf{1}}_n := [\delta_{ij}]_{n\times n}$$

where δ_{ij} *is Kronecker's "delta":*

$$\delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

The matrix of $0 : V \to W, \quad \mathbf{v} \mapsto \mathbf{0}_W$ *with respect to the bases* $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ *for* V *and* $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ *for* W *is*

$$\underline{\mathbf{0}}_{m\times n} := [x_{ij}]_{m\times n}$$

with $x_{ij} = 0$ *for all* i, j.

Proof. Exercise. □

Definition 9.5. $\underline{\mathbf{0}}_{m\times n}$ is the $(m \times n)$ *zero matrix* and $\underline{\mathbf{1}}_n$ the $(n \times n)$ *identity matrix*.

Observation 9.6. The matrix representations of the vector $\mathbf{x} \in V$ and the linear transformation $T : V \to W$ as

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

depend critically upon the choice of bases for V and W, as the next examples illustrate.

Example 9.7. Take $\mathbb{F} := \mathbb{R}$, $V := \mathbb{R}^3$, $W := \mathbb{R}^2$ and $T : V \to W, \ (x, y, z) \mapsto (x, y)$.

Choose $\mathbf{e}_1 := (1, 0, 0)$, $\mathbf{e}_2 := (0, 1, 0)$, $\mathbf{e}_3 := (0, 0, 1) \in V$ and $\mathbf{f}_1 := (1, 0)$, $\mathbf{f}_2 := (0, 1) \in W$.

We verify that the above vectors do, in fact, forms bases for V and W respectively.

Take $\mathbf{x} = (x, y, z) \in V = \mathbb{R}^3$. Then $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ if and only if

$$(x, y, z) = x_1(1, 0, 0) + x_2(0, 1, 0) + x_3(0, 0, 1) = (x_1, x_2, x_3)$$

so that $x_1 = x, x_2 = y, x_3 = z$ is the unique solution. The expression being unique, $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ does form a basis for $\mathbb{R}^3$ and the coordinate vector of $(x, y, z) \in \mathbb{R}^3$ with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

An analogous calculation shows that $\{\mathbf{f}_1, \mathbf{f}_2\}$ is a basis for $W = \mathbb{R}^2$ and that the coordinate vector of $(u, v) \in \mathbb{R}^2$ with respect to the basis $\{\mathbf{f}_1, \mathbf{f}_2\}$ is

$$\begin{bmatrix} u \\ v \end{bmatrix}.$$

Since

$$\begin{aligned} T(\mathbf{e}_1) &= T(1, 0, 0) = (1, 0) &= \mathbf{f}_1 &= 1\mathbf{f}_1 + 0\mathbf{f}_2 \\ T(\mathbf{e}_2) &= T(0, 1, 0) = (0, 1) &= \mathbf{f}_2 &= 0\mathbf{f}_1 + 1\mathbf{f}_2 \\ T(\mathbf{e}_3) &= T(0, 0, 1) = (0, 0) &= \mathbf{0}_W &= 0\mathbf{f}_1 + 0\mathbf{f}_2, \end{aligned}$$

the matrix of T with respect to the bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for V and $\{\mathbf{f}_1, \mathbf{f}_2\}$ for W is

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

The matrix version of $T(x, y, z) = (x, y)$ is thus

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Example 9.8. Take $\mathbb{F} := \mathbb{R}$, $V := \mathbb{R}^3$, $W := \mathbb{R}^2$ and $T : V \to W,\ (x, y, z) \mapsto (x, y)$.

Choose $\mathbf{e}_1 := (1, 0, 0),\ \mathbf{e}_2 := (1, 1, 0),\ \mathbf{e}_3 := (1, 1, 1) \in V$ and $\mathbf{f}_1 := (0, 1), \mathbf{f}_2 := (1, 1) \in W$.

We verify that the above vectors form bases for V and W respectively.

Take $\mathbf{x} = (x, y, z) \in V = \mathbb{R}^3$. Then $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ if and only if

$$(x, y, z) = x_1(1, 0, 0) + x_2(1, 1, 0) + x_3(1, 1, 1) = (x_1 + x_2 + x_3, x_2 + x_3, x_3)$$

so that $x_3 = z, x_2 = y - z, x_1 = x - y$ is the unique solution. Since the expression is unique, $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ forms a basis for $\mathbb{R}^3$, and the coordinate vector of $(x, y, z) \in \mathbb{R}^3$ with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is

$$\begin{bmatrix} x - y \\ y - z \\ z \end{bmatrix}.$$

To see that $\{\mathbf{f}_1, \mathbf{f}_2\}$ is a basis for $W = \mathbb{R}^2$, note that $(u, v) = y_1\mathbf{f}_1 + y_2\mathbf{f}_2$ if and only if

$$(u, v) = y_1(0, 1) + y_2(1, 1) = (y_2, y_1 + y_2),$$

which has the unique solution $y_2 = u, y_1 = v - u$ showing that the coordinate vector of $(u, v) \in \mathbb{R}^2$ with respect to the basis $\{\mathbf{f}_1, \mathbf{f}_2\}$ is

$$\begin{bmatrix} v - u \\ u \end{bmatrix}.$$

Since

$$\begin{aligned} T(\mathbf{e}_1) &= T(1, 0, 0) = (1, 0) &&= -1\mathbf{f}_1 + 1\mathbf{f}_2 \\ T(\mathbf{e}_2) &= T(1, 1, 0) = (1, 1) &&= 0\mathbf{f}_1 + 1\mathbf{f}_2 \\ T(\mathbf{e}_3) &= T(1, 1, 1) = (1, 1) &&= 0\mathbf{f}_1 + 1\mathbf{f}_2, \end{aligned}$$

the matrix of T with respect to the bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for V and $\{\mathbf{f}_1, \mathbf{f}_2\}$ for W is

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

The matrix version of $T(x, y, z) = (x, y)$ is thus

$$\begin{bmatrix} y - x \\ x \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x - y \\ y - z \\ z \end{bmatrix}$$

Example 9.9. Let $\mathbb{F}[t]$ be the set of all polynomials in t with coefficients from $\mathbb{F}$. This is a real vector space with respect to the usual addition of polynomials and the usual multiplication of polynomials by constants:

Recall that as polynomials

$$\sum_{i=0}^{m} a_i t^i = \sum_{j=0}^{n} a_j t^j$$

if and only if $m = n$ and $a_i = b_i$ for each $i \in \{1, \ldots, n\}$.

Plainly, the subset, $\mathcal{P}_n$, of $\mathbb{F}[t]$ consisting of all polynomials, whose degree does not exceed n, forms a vector subspace of $\mathbb{F}[t]$.

Take $\mathbb{F} = \mathbb{R}, V = W = \mathcal{P}_2$ and $T : V \to W, \quad p \mapsto p' = \frac{d}{dt}(p)$, where for $p = a_0 + \cdots + a_n t^n$

$$p' = a_1 + 2a_2 t + \cdots + na_n t^{n-1}.$$

Choose the vectors $\mathbf{e}_1 = \mathbf{f}_1 = 1, \mathbf{e}_2 = \mathbf{f}_2 = t, \mathbf{e}_3 = \mathbf{f}_3 = t^2 \in V(= W)$.

To see that these form a basis of $\mathcal{P}_2$, observe that the definition of a polynomial means that every element of $\mathcal{P}_2$ can be written uniquely as a linear combination of $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$.

It follows that the coordinate vector of $a + bt + ct^2 \in \mathcal{P}_2$ with respect to the basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ is

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Since

$$\begin{aligned} T(\mathbf{e}_1) &= \frac{d}{dt}(1) = 0 &&= 0\mathbf{f}_1 + 0\mathbf{f}_2 + 0\mathbf{f}_3 \\ T(\mathbf{e}_2) &= \frac{d}{dt}(t) = 1 &&= 1\mathbf{f}_1 + 0\mathbf{f}_2 + 0\mathbf{f}_3 \\ T(\mathbf{e}_3) &= \frac{d}{dt}(t^2) = 2t &&= 0\mathbf{f}_1 + 2\mathbf{f}_2 + 0\mathbf{f}_3, \end{aligned}$$

the matrix of T with respect to the bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for V and $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ for W is

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

The matrix version of $T(p) = p'$ with respect to these bases is thus

$$\begin{bmatrix} b \\ 2c \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

Example 9.10. We continue with the vector space V in the previous example.

Choose $\mathbf{e}_1 = t^2, \mathbf{e}_2 = t, \mathbf{e}_3 = 1 \in V$ and $\mathbf{f}_1 = 1, \mathbf{f}_2 = t + 1, \mathbf{f}_3 = t^2 + 1 \in W$.

Plainly $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ forms a basis for V, as it is just a re-ordering of our previous basis.

To see that $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ forms a basis of $\mathcal{P}_2$, note that $p(t) = a + bt + ct^2 = y_1\mathbf{f}_1 + y_2\mathbf{f}_2 + y_3\mathbf{f}_3$ if and only if

$$a + bt + ct^2 = y_1 + y_2(t + 1) + y_3(t^2 + 1) = (y_1 + y_2 + y_3) + y_2 t + y_3 t^2,$$

which is the case if and only if $y_1 = a - b - c, y_2 = b, y_3 = c$.

The coordinate vector of $a + bt + ct^2 \in V$ with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is therefore

$$\begin{bmatrix} c \\ b \\ a \end{bmatrix},$$

and that of $a + bt + ct^2 \in W$ with respect to the basis $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ is

$$\begin{bmatrix} a - b - c \\ b \\ c \end{bmatrix}$$

Since

$$\begin{aligned} T(\mathbf{e}_1) &= \frac{d}{dx}(x^2) = 2x = -2\mathbf{f}_1 + 2\mathbf{f}_2 + 0\mathbf{f}_3 \\ T(\mathbf{e}_2) &= \frac{d}{dx}(x) = 1 = 1\mathbf{f}_1 + 0\mathbf{f}_2 + 0\mathbf{f}_3 \\ T(\mathbf{e}_3) &= \frac{d}{dx}(1) = 0 = 0\mathbf{f}_1 + 0\mathbf{f}_2 + 0\mathbf{f}_3, \end{aligned}$$

the matrix of T with respect to the bases $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ for V and $\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\}$ for W is

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

The matrix version of $T(p) = p'$ with respect to these bases is thus

$$\begin{bmatrix} b - 2c \\ 2c \\ 0 \end{bmatrix} = \begin{bmatrix} -2 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} c \\ b \\ a \end{bmatrix}$$

9.1 The Matrix of a Composite Linear Transformation

Now that we have seen how to use matrices to represent linear transformations between finitely generated vector spaces, we consider the relationship between matrices representing two linear transformations and the matrix representing their composition. (Of course, we assume that the linear transformations in question can be composed and, in order to be able to use matrices, that we have chosen a fixed basis for each of the vector spaces in question.)

We begin with a concrete example.

Example 9.11. We consider the linear transformations

$$\begin{aligned} &S\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad (x, y) \longmapsto (x + 2y, 3x - y) \\ &R\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2, \quad (u, v) \longmapsto (2u + 4v, -5u + 7v) \end{aligned}$$

We use the basis $\{(1, 0), (0, 1)\}$ for each of the three vector spaces.

Let $\underline{\mathbf{A}}$ be the matrix of R and $\underline{\mathbf{B}}$ that of S with respect to the chosen bases. Then

$$\underline{\mathbf{A}} = \begin{bmatrix} 2 & 4 \\ -5 & 7 \end{bmatrix} \quad \text{and} \quad \underline{\mathbf{B}} = \begin{bmatrix} 1 & 2 \\ 3 & -1 \end{bmatrix}.$$

The composition of these linear transformations, $T = R \circ S : \mathbb{R}^2 \to \mathbb{R}^2$ is given by

$$\begin{aligned} (R \circ S)(x, y) &:= R(S(x, y)) \\ &:= R(x + 2y, 3x - y) \\ &:= (2(x + 2y) + 4(3x - y), -5(x + 2y) + 7(3x - y)) \\ &:= ((2.1 + 4.3)x + (2.2 + 4(-1))y, ((-5).1 + 7.3)x + ((-5).2 + 7.(-1))y) \\ &:= (14x, 16x - 17y) \end{aligned}$$

Let $\underline{\mathbf{C}}$ be the matrix of T with respect to the chosen bases, so that

$$\underline{\mathbf{C}} = \begin{bmatrix} 14 & 0 \\ 16 & -17 \end{bmatrix} = \begin{bmatrix} 2.1+4.3 & 2.2+4(-1) \\ (-5).1+7.3 & (-5).2+7.(-1) \end{bmatrix}$$

The intermediate steps have been included because they provide the pattern for the general solution. We formulate the above argument and calculation in sufficiently general terms.

Take finitely generated vector spaces U, V, W over the field $\mathbb{F}$, and linear transformations $S : U \to V$ and $R : V \to W$. Let $T = R \circ S \colon U \to W$ be their composition.

Choose bases $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ for U, $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ for V and $\{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ for W.

If $\mathbf{x} \in U, \mathbf{y} \in V$ and $\mathbf{z} \in W$ there are unique $x_\ell, y_k, z_j \in \mathbb{F}$ $(1 \leq \ell \leq p, 1 \leq k \leq n, 1 \leq j \leq m)$ with

$$\mathbf{x} = \sum_{\ell=1}^{p} x_\ell \mathbf{u}_\ell, \quad \mathbf{y} = \sum_{k=1}^{n} y_k \mathbf{v}_k, \quad \text{and} \quad \mathbf{z} = \sum_{j=1}^{m} z_j \mathbf{w}_j,$$

and, by Theorem 5.3, there are unique $a_{ij}, b_{kl} \in \mathbb{F}$ $(1 \leq i \leq m,\ 1 \leq j, k \leq n,\ 1 \leq \ell \leq p)$ with

$$S(\sum_{l=1}^{p} x_l \mathbf{u}_l) = \sum_{k=1}^{n} \left(\sum_{l=1}^{p} b_{kl} x_l \right) \mathbf{v}_k \tag{9.7}$$

$$R(\sum_{k=1}^{n} y_k \mathbf{v}_k) = \sum_{j=1}^{m} \left(\sum_{k=1}^{n} a_{jk} y_k \right) \mathbf{w}_j \tag{9.8}$$

In other words, the matrices of S and R with respect to the chosen bases are, respectively,

$$\underline{\mathbf{B}} := \begin{bmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & & \vdots \\ b_{m1} & \cdots & b_{np} \end{bmatrix} \quad \text{and} \quad \underline{\mathbf{A}} := \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{mn} \end{bmatrix}.$$

Thus $S(\mathbf{x}) = \mathbf{y}$ becomes

$$S(\sum_{l=1}^{p} x_l \mathbf{u}_l) = \sum_{k=1}^{n} y_k \mathbf{v}_k,$$

whence

$$y_k = \sum_{l=1}^{p} b_{kl} x_l.$$

Similarly, $R(\mathbf{y}) = \mathbf{z}$ becomes

$$R(\sum_{k=1}^{n} y_k \mathbf{v}_k) = \sum_{j=1}^{m} z_j \mathbf{w}_j,$$

whence

$$z_j = \sum_{k=1}^{n} a_{jk} y_k.$$

Finally, $T(\mathbf{x}) = \mathbf{z}$ becomes

$$T(\sum_{\ell=1}^{p} x_k \mathbf{u}_\ell) = \sum_{j=1}^{m} z_j \mathbf{w}_j,$$

whence

$$z_j = \sum_{\ell=1}^{p} c_{j\ell} x_\ell.$$

But, $T = R \circ S$, so that $T(\mathbf{x}) = R(S(\mathbf{x})) = R(\mathbf{y})$. Then, by direct substitution,

$$\begin{aligned} z_j &= \sum_{k=1}^{n} a_{jk} y_k && \text{as } R(\mathbf{y}) = \mathbf{z} \\ &= \sum_{k=1}^{n} a_{jk} \left(\sum_{l=1}^{p} b_{kl} x_l \right) && \text{as } S(\mathbf{x}) = \mathbf{y} \\ &= \sum_{k=1}^{n} \left(\sum_{l=1}^{p} a_{jk} b_{kl} x_l \right) \\ &= \sum_{l=1}^{p} \left(\sum_{k=1}^{n} a_{jk} b_{kl} \right) x_l. \end{aligned}$$

Thus.

$$c_{jl} := \sum_{k=1}^{n} a_{jk} b_{kl}, \qquad j = 1, \dots m, \ l = 1, \dots, p$$

Observation 9.12. We began with the aim of finding a way of calculating the matrix of the composition of two linear transformations directly from the matrices representing them. Our calculations, which, when examined closely, consist simply of substitution of variables, show that there is one and only possible way of doing so.

This not only derives the formula familiar from prior exposure to matrices, but also explains why matrix multiplication is defined only when the number of columns in the left-hand matrix agrees with the number of rows in the right-hand matrix.

This illustrates how purely abstract considerations can lead to explicit and convenient computational methods.

Definition 9.13. Let $[a_{ij}]_{m \times n}$ and $[b_{jk}]_{n \times p}$ be matrices over the field $\mathbb{F}$. Their *product* is the matrix $[c_{ik}]_{m \times p}$ with

$$c_{jl} := \sum_{k=1}^{n} a_{jk} b_{kl}, \tag{9.9}$$

for all $j = 1, \dots m, \ l = 1, \dots, p$.

The next theorem summarises the above.

Theorem 9.14. *Take two composable linear transformation. Choose a fixed basis for each of the vector spaces concerned, and take matrices of all linear transformations with respect to these bases. Then the matrix of the composite linear transformation is the product of the matrices of the linear transformations.*

Corollary 9.15. *Matrix multiplication is associative.*

Proof. The composition of functions, in particular, linear transformations, is associative. $\square$

Corollary 9.16. *For any $m \times n$ matrix $\underline{\mathbf{A}}$,*

$$\underline{\mathbf{1}}_m \underline{\mathbf{A}} = \underline{\mathbf{A}} = \underline{\mathbf{A}\mathbf{1}}_n$$

Proof. The assertions follow directly from the definition of matrix multiplication. □

Definition 9.17. In *inverse* of the $m \times n$ matrix $\underline{\mathbf{A}}$ is an $n \times m$ matrix $\underline{\mathbf{B}}$ such that

$$\underline{\mathbf{A}}\,\underline{\mathbf{B}} = \underline{\mathbf{1}}_m \quad \text{and} \quad \underline{\mathbf{B}}\,\underline{\mathbf{A}} = \underline{\mathbf{1}}_n$$

In such a case, $\underline{\mathbf{A}}$ is said to be *invertible*

Theorem 9.18. *The $m \times n$ matrix $\underline{\mathbf{A}}$ is invertible only if $m = n$. If $\underline{\mathbf{A}}$ is invertible, then its inverse is uniquely determined.*

Proof. Exercise. □

9.2 The Matrix of the Sum of Linear Transformations

The matrix representation of the composition of linear transformation led to the definition of matrix multiplication.

It is natural to seek matrix versions of the other operations on linear transformations we have met.

Definition 6.21 introduced addition of linear transformations $S, T \colon V \to W$ by

$$(S \boxplus T)(\mathbf{v}) := S(\mathbf{v}) + T(\mathbf{v}) \qquad (\mathbf{v} \in V).$$

Choose bases, $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ for V and $\{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ for W. Let $[a_{ij}]_{m\times n}$ and $[b_{ij}]_{m\times n}$ be the respective matrices of S and T with respect to the chosen bases, and suppose $\mathbf{v} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$. Then

$$S(\mathbf{v}) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)\mathbf{f}_i$$

$$T(\mathbf{v}) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)\mathbf{f}_i$$

and so, by definition,

$$\begin{aligned}(S \boxplus T)(\mathbf{v}) &= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)\mathbf{f}_i + \sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)\mathbf{f}_i \\ &= \sum_{i=1}^{m}\left(\sum_{j=1}^{n} (a_{ij} + b_{ij})x_j\right)\mathbf{f}_i\end{aligned}$$

If $[c_{ij}]_{m\times n}$ is the matrix pf $S \boxplus T$ with respect to the given bases, then

$$(S \boxplus T)(\mathbf{v}) = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} c_{ij}x_j\right)\mathbf{f}_j,$$

whence, by uniqueness,

$$c_{ij} = a_{ij} + b_{ij}$$

This motivates our definition.

Definition 9.19. Given $m \times n$ matrices over $\mathbb{F}$, $[a_{ij}]_{m\times n}, [b_{ij}]_{m\times n}$, their *sum* is the $m \times n$ matrix, $['c_{ij}]_{m\times n}$, where $c_{ij} := a_{ij} + b_{ij}$.

Lemma 9.20. *Addition of matrices is associative and commutative, has the zero matrix as neutral element and the additive inverse of* $[a_{ij}]_{m\times n}$ *is* $[-a_{ij}]_{m\times n}$

Proof. Immediate from the corresponding properties of linear transformations. □

Observation 9.21. Note that an $m \times n$ matrix and a $p \times q$ matrix can be added if and only if $p = m$ and $q = n$. This is a consequence of the fact that the sum of the linear transformations S and T is defined when and only when $\mathrm{dom}(S) = \mathrm{dom}(T)$ and $\mathrm{codom}(S) = \mathrm{codom}(T)$.

Moreover, in this case, we can also multiply matrices by scalars, based on scalar multiplication of functions.

Definition 9.22. Given $\lambda \in \mathbb{F}$ and $\underline{\mathbf{A}} = [a_{ij}]_{m\times n}$ over $\mathbb{F}$, $\lambda\underline{\mathbf{A}} = [c_{ij}]_{m\times n}$. where $c_{ij} := \lambda a_{ij}$.

9.3 Algebraic Properties of the Operations on Matrices

Since the operations we have introduced on matrices represent operations we defined on linear transformations in Chapter 6, we can translate the properties of the operations on linear transformations into properties of operations of matrices.

We summarise these in a theorem whose proof is left to the reader.

Theorem 9.23. *Take* $\underline{\mathbf{A}}, \underline{\mathbf{B}} \in \mathbf{M}(m \times n; \mathbb{F})$ *and* $\underline{\mathbf{C}}, \underline{\mathbf{D}} \in \mathbf{M}(n \times p; \mathbb{F})$.

(i) $\underline{\mathbf{1}}_m \underline{\mathbf{A}} = \underline{\mathbf{A}} = \underline{\mathbf{A}}\,\underline{\mathbf{1}}_n$

(ii) $\underline{\mathbf{A}}(\underline{\mathbf{C}} + \underline{\mathbf{D}}) = (\underline{\mathbf{A}}\,\underline{\mathbf{C}}) + (\underline{\mathbf{A}}\,\underline{\mathbf{D}})$

(iii) $(\underline{\mathbf{A}} + \underline{\mathbf{B}})\underline{\mathbf{C}} = (\underline{\mathbf{A}}\,\underline{\mathbf{C}}) + (\underline{\mathbf{B}}\,\underline{\mathbf{C}})$

(iv) $L_{\underline{\mathbf{A}}} \colon \mathbf{M}(n \times p; \mathbb{F}) \longrightarrow \mathbf{M}(m \times p; \mathbb{F}), \quad \underline{\mathbf{C}} \longmapsto \underline{\mathbf{A}}\,\underline{\mathbf{C}}$ *is a linear transformation.*

(v) $R_{\underline{\mathbf{C}}} \colon \mathbf{M}(m \times n; \mathbb{F}) \longrightarrow \mathbf{M}(m \times p; \mathbb{F}), \quad \underline{\mathbf{A}} \longmapsto \underline{\mathbf{A}}\,\underline{\mathbf{C}}$ *is a linear transformation.*

9.4 Matrices Representing the Same Linear Transformation

In order to represent a linear transformation between finite dimensional vector spaces by a matrix, we need to choose a basis for each of the vector spaces in question, and the resulting matrix depends on the particular bases chosen. It is, therefore, natural to ask:

> What is the relationship between two matrices representing one and the same linear transformation between two given finitely generated vector spaces?

This section is devoted to this question.

We know that each finitely generated vector space is determined up to isomorphism by a single (numerical) invariant, its dimension, which is the number of vectors in any basis for it.

Moreover, we have also seen that if $\dim_{\mathbb{F}} V = n$ and $\dim_{\mathbb{F}} W = m$, then any matrix representing the linear transformation $T \colon V \to W$ must be an $m \times n$ matrix over $\mathbb{F}$. Thus, one immediate necessary condition for two matrices to represent one and the same linear transformation between two finite dimensional vector spaces is that they both be "of the same size".

The following example shows that this necessary condition is not sufficient.

Example 9.24. No linear transformation can be represented by both $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

One way to see why these matrices cannot represent the same linear transformation is to observe that any linear transformation represented by the former must be bijective, whereas no linear transformation represented by the latter can be either injective or surjective. (We leave it to the reader to contemplate why these statements are true. The reasons will become evident later.)

Consider the linear transformation $T\colon V \to W$ from the n-dimensional vector space V to the m-dimensional vector space W.

Let $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ and $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ be bases for V and $\{\mathbf{f}_1, \dots, \mathbf{f}_m\}$ and $\{\mathbf{f}'_1, \dots, \mathbf{f}'_m\}$ for W.

Let the matrix of T with respect to $\{\mathbf{e}_j\}$ and $\{\mathbf{f}_i\}$ be $\underline{\mathbf{A}} = [a_{ij}]_{m\times n}$, and its matrix with respect to $\{\mathbf{e}'_j\}$ and $\{\mathbf{f}'_i\}$ be $\underline{\mathbf{A}}' = [a'_{ij}]_{m\times n}$. Then

$$T(\mathbf{e}_j) = \sum_{i=1}^{m} a_{ij}\mathbf{f}_i = a_{1j}\mathbf{f}_1 + \cdots + a_{mj}\mathbf{f}_m \quad \text{and} \quad T(\mathbf{e}'_j) = \sum_{i=1}^{m} a'_{ij}\mathbf{f}'_i = a'_{1j}\mathbf{f}'_1 + \cdots + a'_{mj}\mathbf{f}'_m.$$

Finally, let

$$\mathbf{e}'_\ell = \sum_{j=1}^{n} \lambda_{j\ell}\mathbf{e}_j = \lambda_{1\ell}\mathbf{e}_1 + \cdots + \lambda_{n\ell}\mathbf{e}_n \quad \text{and} \quad \mathbf{f}'_k = \sum_{i=1}^{m} \mu_{ik}\mathbf{f}_i = \mu_{1k}\mathbf{f}_1 + \cdots + \mu_{mk}\mathbf{f}_m.$$

In other words, the co-ordinate vector of $\mathbf{e}'_\ell$ with respect to the basis $\{\mathbf{e}_j \mid j = 1, \dots, n\}$ and that of $\mathbf{f}_k$ with respect to $\{\mathbf{f}_i \mid i = 1, \dots, m\}$ are, respectively,

$$\begin{bmatrix} \lambda_{1\ell} \\ \vdots \\ \lambda_{n\ell} \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} \mu_{1k} \\ \vdots \\ \mu_{mk} \end{bmatrix}$$

We form the $n \times n$ matrix $\underline{\mathbf{L}} := [\lambda_{ij}]_{n\times n}$ and the $m \times m$ matrix $\underline{\mathbf{M}} := [\mu_{ij}]_{m\times m}$.

Take $\mathbf{v} \in V$ and let $T(\mathbf{v}) = \mathbf{w} \in W$. Then

$$\begin{aligned} \mathbf{v} &= \sum_{\ell=1}^{n} x'_\ell \mathbf{e}'_\ell \\ &= \sum_{\ell=1}^{n} x'_\ell \left(\sum_{j=1}^{n} \lambda_{j\ell}\mathbf{e}_j \right) \\ &= \sum_{j=1}^{n} \left(\sum_{\ell=1}^{n} \lambda_{j\ell} x'_\ell \right) \mathbf{e}_j \end{aligned}$$

But $\mathbf{v} = \sum_{j=1}^{n} x_j e_j$, with the x_j's unique. So $x_j = \sum_{\ell=1}^{n} \lambda_{j\ell} x'_\ell$, or

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \lambda_{11} & \cdots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{n1} & \cdots & \lambda_{nn} \end{bmatrix} \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix}.$$

Similarly, if $\mathbf{w} \in W$, $\mathbf{w} = \sum_{k=1}^{m} y'_k \mathbf{e}'_k = \sum_{k=1}^{m} y_k \mathbf{e}_k$, with

$$\begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} = \begin{bmatrix} \mu_{11} & \dots & \mu_{1m} \\ \vdots & & \vdots \\ \mu_{m1} & \dots & \mu_{mm} \end{bmatrix} \begin{bmatrix} y'_1 \\ \vdots \\ y'_m \end{bmatrix}$$

Consequently,

$$\begin{aligned} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} &= \begin{bmatrix} \mu_{11} & \dots & \mu_{1m} \\ \vdots & & \vdots \\ \mu_{m1} & \dots & \mu_{mm} \end{bmatrix} \begin{bmatrix} y'_1 \\ \vdots \\ y'_m \end{bmatrix} \\ &= \begin{bmatrix} \mu_{11} & \dots & \mu_{1m} \\ \vdots & & \vdots \\ \mu_{m1} & \dots & \mu_{mm} \end{bmatrix} \begin{bmatrix} a'_{11} & \dots & a'_{1n} \\ \vdots & & \vdots \\ a'_{m1} & \dots & a'_{mn} \end{bmatrix} \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix} \end{aligned}$$

On the other hand,

$$\begin{aligned} \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} &= \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} \lambda_{11} & \dots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{n1} & \dots & \lambda_{nn} \end{bmatrix} \begin{bmatrix} x'_1 \\ \vdots \\ x'_n \end{bmatrix} \end{aligned}$$

Since these equations hold for all $x'_1, \dots, x'_n \in \mathbb{F}$, it follows that

$$\begin{bmatrix} \mu_{11} & \dots & \mu_{1m} \\ \vdots & & \vdots \\ \mu_{m1} & \dots & \mu_{mm} \end{bmatrix} \begin{bmatrix} a'_{11} & \dots & a'_{1n} \\ \vdots & & \vdots \\ a'_{m1} & \dots & a'_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} \lambda_{11} & \dots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{n1} & \dots & \lambda_{nn} \end{bmatrix}.$$

This leads us to the next definition and theorem, summarising the above.

Definition 9.25. The matrix $\underline{\mathbf{L}} := \begin{bmatrix} \lambda_{11} & \dots & \lambda_{1n} \\ \vdots & & \vdots \\ \lambda_{n1} & \dots & \lambda_{nn} \end{bmatrix}$ is called a *change of basis matrix (from the basis $\{\mathbf{e}'_i\}$ to the basis $\{\mathbf{e}_i\}$).*

Of course, $\underline{\mathbf{M}} := \begin{bmatrix} \mu_{11} & \dots & \mu_{1m} \\ \vdots & & \vdots \\ \mu_{m1} & \dots & \mu_{mm} \end{bmatrix}$ is also a change of basis matrix.

Theorem 9.26. *If $\underline{\mathbf{A}}$ and $\underline{\mathbf{A}}'$ are the matrices of the linear transformation $T\colon V \to W$ and if $\underline{\mathbf{L}}$ and $\underline{\mathbf{M}}$ are the change of basis matrices from the bases which give rise to $\underline{\mathbf{A}}'$ to the bases which give rise to $\underline{\mathbf{A}}$, then*

$$\underline{\mathbf{M}}\,\underline{\mathbf{A}}' = \underline{\mathbf{A}}\,\underline{\mathbf{L}}.$$

Corollary 9.27. *If $\underline{\mathbf{A}}$ and $\underline{\mathbf{B}}$ are the matrices of the linear transformation $T\colon V \to V$ and if $\underline{\mathbf{M}}$ is the change of basis matrix from the basis which gives rise to $\underline{\mathbf{B}}$ to the basis which give rise to $\underline{\mathbf{A}}$, then, for any counting number n,*

$$\underline{\mathbf{B}}^n = \underline{\mathbf{M}}^{-1}\underline{\mathbf{A}}^n\,\underline{\mathbf{M}}.$$

Proof. By Theorem 9.26, $\underline{\mathbf{M}}\,\underline{\mathbf{B}} = \underline{\mathbf{A}}\,\underline{\mathbf{M}}$, or, eqivalently, $\underline{\mathbf{B}} = \underline{\mathbf{M}}^{-1}\underline{\mathbf{A}}\,\underline{\mathbf{M}}$. Thus,

$$\begin{aligned}
\underline{\mathbf{B}}^n &= (\underline{\mathbf{M}}^{-1}\underline{\mathbf{A}}\,\underline{\mathbf{M}})^n \\
&= (\underline{\mathbf{M}}^{-1}\underline{\mathbf{A}}\,\underline{\mathbf{M}})(\underline{\mathbf{M}}^{-1}\underline{\mathbf{A}}\,\underline{\mathbf{M}})\cdots(\underline{\mathbf{M}}^{-1}\underline{\mathbf{A}}\,\underline{\mathbf{M}}) \\
&= \underline{\mathbf{M}}^{-1}\underline{\mathbf{A}}\,(\underline{\mathbf{M}}\,\underline{\mathbf{M}}^{-1})\underline{\mathbf{A}}\,(\underline{\mathbf{M}}\,\underline{\mathbf{M}}^{-1})\cdots(\underline{\mathbf{M}}\,\underline{\mathbf{M}}^{-1})\underline{\mathbf{A}}\,\underline{\mathbf{M}} && \text{by associativity} \\
&= \underline{\mathbf{M}}^{-1}\underline{\mathbf{A}}^n\underline{\mathbf{M}} && \text{since } \underline{\mathbf{M}}\,\underline{\mathbf{M}}^{-1} = \underline{\mathbf{1}}_n.
\end{aligned}$$

□

Observation 9.28. Examples 2.4, 2.5, 2.6 and 2.8 are applications of this corollary.

Example 9.29. Let $\begin{bmatrix} 4 & -3 \\ 1 & 0 \end{bmatrix}$ to be the matrix of a linear transformation $T\colon V \to V$ with respect to the basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ of the real vector space V. Thus $\dim_{\mathbb{R}}(V) = 2$.
Put $\mathbf{e}_1' := 3\mathbf{e}_1 + \mathbf{e}_2$ and $\mathbf{e}_2' := \mathbf{e}_1 + \mathbf{e}_2$.
By direct calculation, $\mathbf{e}_1 = \frac{1}{2}\left(\mathbf{e}_1' - \mathbf{e}_2'\right)$ and $\mathbf{e}_2 = \frac{1}{2}\left(-\mathbf{e}_1' + 3\mathbf{e}_2'\right)$.
Thus $\mathbf{e}_1, \mathbf{e}_2 \in \langle \mathbf{e}_1', \mathbf{e}_2' \rangle$. It follows that

$$V = \langle \mathbf{e}_1, \mathbf{e}_2 \rangle \subseteq \langle \mathbf{e}_1', \mathbf{e}_2' \rangle \subseteq V$$

Since they generate the two-dimensional vector space V, the vectors $\mathbf{e}_1'$ and $\mathbf{e}_2'$ form a basis for V.
Let $\mathbf{v} \in V$ have co-ordinate vector $\begin{bmatrix} r \\ s \end{bmatrix}$ with respect to $\{\mathbf{e}_1, \mathbf{e}_2\}$ and $\begin{bmatrix} x \\ y \end{bmatrix}$ with respect to $\{\mathbf{e}_1', \mathbf{e}_2'\}$.
As $x\mathbf{e}_1' + y\mathbf{e}_2' = (3x + y)\mathbf{e}_1 + (x + y)\mathbf{e}_2$,

$$\begin{bmatrix} r \\ s \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix},$$

and since $r\mathbf{e}_1 + s\mathbf{e}_2 = \frac{1}{2}\Big((r - s)\mathbf{e}_1' + (-r + 3s)\mathbf{e}_2'\Big)$,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} r \\ s \end{bmatrix},$$

It follows that the matrix of T with respect to the basis $\{\mathbf{e}_1', \mathbf{e}_2'\}$ is

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

Observation 9.30. Every $n \times n$ matrix with coefficients in $\mathbb{F}$ is the matrix of a linear transformation $T : V \to V$. It is natural to ask:

> Which linear transformation is represented by the change of basis matrix $\underline{\mathbf{L}}$?

To answer this, note that when we change the basis, we do *not* change V: Each vector $\mathbf{v} \in V$ is left unchanged, only the co-ordinate vector we assign to it changes. In other words, $\underline{\mathbf{L}}$ is the matrix of the identity linear transformation $id_V\colon V \to V,\ \mathbf{v} \mapsto \mathbf{v}$.

9.5 Another Look at Matrix Multiplication

We investigate the relationship between the rows (resp. columns) of the product of two matrices and the rows (resp. columns) of the matrices being multiplied.

Take $\underline{\mathbf{A}} = [a_{ij}]_{m\times n}, \underline{\mathbf{B}} = [b_{jk}]_{n\times p}$ and $\underline{\mathbf{C}} = [c_{ik}]_{m\times p}$ and suppose that $\underline{\mathbf{C}} = \underline{\mathbf{A}}\,\underline{\mathbf{B}}$, that is

$$c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk} \qquad (1 \le i \le m,\ 1 \le k \le p)$$

The k^{th} column of $\underline{\mathbf{C}}$ is obtained by fixing k, in which case we obtain

$$\begin{aligned}
\begin{bmatrix} c_{1k} \\ \vdots \\ c_{mk} \end{bmatrix} &= \begin{bmatrix} a_{11}b_{1k} + \cdots + a_{1n}b_{nk} \\ \vdots \\ a_{m1}b_{1k} + \cdots + a_{mn}b_{nk} \end{bmatrix} \\
&= \begin{bmatrix} a_{11}b_{1k} \\ \vdots \\ a_{m1}b_{1k} \end{bmatrix} + \cdots + \begin{bmatrix} a_{1n}b_{nk} \\ \vdots \\ a_{mn}b_{nk} \end{bmatrix} \\
&= \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix} b_{1k} + \cdots + \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix} b_{nk}
\end{aligned}$$

In other words, the k^{th} column of $\underline{\mathbf{A}}\,\underline{\mathbf{B}}$ is the linear combination of the columns of $\underline{\mathbf{A}}$ given by the entries in the k^{th} column of $\underline{\mathbf{B}}$.

Now fix i instead of k.

$$\begin{aligned}
\begin{bmatrix} c_{i1} & \cdots & c_{ip} \end{bmatrix} &= \begin{bmatrix} a_{i1}b_{11} + \cdots + a_{in}b_{n1} & \cdots & a_{i1}b_{1p} + \cdots + a_{in}b_{np} \end{bmatrix} \\
&= \begin{bmatrix} a_{i1}b_{11} & \cdots & a_{i1}b_{1p} \end{bmatrix} + \cdots + \begin{bmatrix} a_{in}b_{n1} & \cdots & a_{in}b_{np} \end{bmatrix} \\
&= a_{i1}\begin{bmatrix} b_{11} & \cdots & b_{1p} \end{bmatrix} + \cdots + a_{in}\begin{bmatrix} b_{n1} & \cdots & b_{np} \end{bmatrix}
\end{aligned}$$

In other words, the i^{th} row of $\underline{\mathbf{A}}\,\underline{\mathbf{B}}$ is the linear combination of the rows of $\underline{\mathbf{B}}$ given by the entries in the i^{th} row of $\underline{\mathbf{A}}$.

Observation 9.31. An important consequence is that such operations on the rows (resp. columns) of a matrix as the elementary row (resp. column) operations can be performed by multiplying the given matrix on the left (resp. right) by a suitable matrix. We return to this later.

9.6 Exercises

Exercise 9.1. Prove that if the linear transformation $T : V \to W$ has matrix $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ with respect to some bases, then it is neither injective nor surjective.

Exercise 9.2. Prove that if linear transformation $T : V \to W$ has matrix $\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ with respect to some bases, then it has a right inverse, but no left inverse.

Exercise 9.3. Let $T : \mathbb{R}^3 \to \mathbb{R}^3$ be rotation about the y-axis through an angle of θ. Show that T is a linear transformation of $\mathbb{R}^3$ to itself and find a matrix representation of T with respect to the bases

(a) $\{(1,0,0),(0,1,0),(0,0,1)\}$ for both the domain and codomain;

(b) $\{(1,0,0),(0,1,0),(0,0,1)\}$ for the domain and $\{(1,0,0),(1,1,0),(1,1,1)\}$ for the codomain;

(c) $\{(1,0,0),(1,1,0),(1,1,1)\}$ for both the domain and codomain;

(d) $\{(1,0,1),(0,1,0),(-1,0,1)\}$ for both the domain and codomain.

Exercise 9.4. Show that the linear transformation $T : V \to W$ is an isomorphism if and only if every matrix which represents it is invertible.

Exercise 9.5. Use the definition of the multiplication of a linear transformation by a scalar to define a multiplication of matrices by scalars.

Exercise 9.6. Prove that if the $m \times n$ matrix $\underline{\mathbf{A}}$ is invertible, then $m = n$ and its inverse is uniquely determined.

Exercise 9.7. Let $\{\mathbf{e}_1, \mathbf{e}_2$ be a basis for the vector space V and $T\colon V \to V$ a linear transformation. Show that in each of the following cases $\{\mathbf{e}'_1, \mathbf{e}'_2\}\}$ is also a basis for V and find the matrix $\underline{\mathbf{A}}'$ of T with respect to the basis $\{\mathbf{e}'_1, \mathbf{e}'_2\}$ given that its matrix with respect to $\{\mathbf{e}_1, \mathbf{e}_2\}$ is $\underline{\mathbf{A}}$.

(a) $\mathbf{e}'_1 := 2\mathbf{e}_1 + \mathbf{e}_2,\ \ \mathbf{e}'_2 := \mathbf{e}_1$ and $\underline{\mathbf{A}} := \begin{bmatrix} 4 & -4 \\ 1 & 0 \end{bmatrix}$

(b) $\mathbf{e}'_1 := (2\sqrt{2}+1)\mathbf{e}_1 + \mathbf{e}_2,\ \ \mathbf{e}'_2 := \mathbf{e}_1 - (\sqrt{2}+1)\mathbf{e}_2$ and $\underline{\mathbf{A}} := \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix}$.