LINEAR TRANSFORMATIONS AND BASES

Let A be a linear transformation from $\mathcal{V}_1$ to $\mathcal{V}_2$, and let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be a basis for $\mathcal{V}_1$. Then if

$$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n$$

is any vector in $\mathcal{V}_1$,

$$A(\mathbf{x}) = x_1A(\mathbf{e}_1) + \cdots + x_nA(\mathbf{e}_n), \tag{2–22}$$

and it follows that the value of $A(\mathbf{x})$ is completely determined by the vectors $A(\mathbf{e}_1), \ldots, A(\mathbf{e}_n)$ in $\mathcal{V}_2$; i.e., A is uniquely determined by its values on a basis for $\mathcal{V}_1$. Moreover, if $\mathbf{y}_1, \ldots, \mathbf{y}_n$ are *arbitrary* vectors in $\mathcal{V}_2$, the mapping $A: \mathcal{V}_1 \to \mathcal{V}_2$ defined by setting

$$A(\mathbf{e}_1) = \mathbf{y}_1, \ \ldots \ , A(\mathbf{e}_n) = \mathbf{y}_n,$$

and then using (2–22) to compute the value of $A(\mathbf{x})$ for every $\mathbf{x}$ in $\mathcal{V}_1$ is clearly linear. Thus Eq. (2–22) also tells us how to construct *all* linear transformations from $\mathcal{V}_1$ to $\mathcal{V}_2$, and we have proved

Theorem 2–5. *Every linear transformation from $\mathcal{V}_1$ to $\mathcal{V}_2$ is uniquely determined by its values on a basis for $\mathcal{V}_1$. These values can be chosen arbitrarily in $\mathcal{V}_2$, different choices yielding different transformations, and every linear transformation from $\mathcal{V}_1$ to $\mathcal{V}_2$ can be obtained in this way.*

EXAMPLE 1. Let $\mathbf{e}_1$ and $\mathbf{e}_2$ be the standard basis vectors in $\mathcal{R}^2$, and let A be a linear transformation from $\mathcal{R}^2$ to $\mathcal{R}^1$. Then A is completely determined by the pair of real numbers $A(\mathbf{e}_1)$, $A(\mathbf{e}_2)$, and can therefore be represented by the ordered pair $\big(A(\mathbf{e}_1), A(\mathbf{e}_2)\big)$. Since distinct ordered pairs define distinct linear transformations, it follows that there are exactly as many linear transformations from $\mathcal{R}^2$ to $\mathcal{R}^1$ as there are vectors in $\mathcal{R}^2$.

EXAMPLE 2. Let $A\colon \mathcal{R}^2 \to \mathcal{R}^2$ be the linear transformation defined by

$$\begin{aligned} A(\mathbf{e}_1) &= (\alpha_1, \alpha_2), \\ A(\mathbf{e}_2) &= (\beta_1, \beta_2), \end{aligned}$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are the standard basis vectors, and let

$$\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2$$

be any vector in $\mathcal{R}^2$. Then

$$\begin{aligned} A(\mathbf{x}) &= x_1A(\mathbf{e}_1) + x_2A(\mathbf{e}_2) \\ &= x_1(\alpha_1, \alpha_2) + x_2(\beta_1, \beta_2) \\ &= (\alpha_1x_1 + \beta_1x_2,\ \alpha_2x_1 + \beta_2x_2). \end{aligned}$$

In this case A can be represented by the array of scalars

$$\begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix},$$

and every such array can be viewed as the definition of a linear transformation A of $\mathcal{R}^2$ into itself where

$$A(x_1, x_2) = (\alpha_1x_1 + \beta_1x_2, \alpha_2x_1 + \beta_2x_2).$$

EXAMPLE 3. Let $A\colon \mathcal{R}^2 \to \mathcal{R}^1$ be the linear transformation described by the ordered pair of real numbers $(2, -1)$ with respect to the standard basis in $\mathcal{R}^2$; that is,

$$A(\mathbf{e}_1) = 2, \qquad A(\mathbf{e}_2) = -1,$$

and let $\mathbf{e}_1' = \mathbf{e}_1 + \mathbf{e}_2$, $\mathbf{e}_2' = -\mathbf{e}_1$. Then $\mathbf{e}_1'$, $\mathbf{e}_2'$ is also a basis for $\mathcal{R}^2$, and we have

$$\begin{aligned} A(\mathbf{e}_1') &= A(\mathbf{e}_1) + A(\mathbf{e}_2) = 1, \\ A(\mathbf{e}_2') &= -A(\mathbf{e}_1) = -2. \end{aligned}$$

Thus the ordered pair which describes A with respect to the basis $\mathbf{e}_1'$, $\mathbf{e}_2'$ is $(1, -2)$, and we see that *the description of a linear transformation by means of its values on a basis changes with a change of basis.*

We have seen that *every* linear transformation $A\colon \mathcal{V}_1 \to \mathcal{V}_2$ can be obtained from the formula

$$A(\mathbf{x}) = x_1A(\mathbf{e}_1) + \cdots + x_nA(\mathbf{e}_n) \tag{2-23}$$

by suitably choosing the $A(\mathbf{e}_j)$ in $\mathcal{V}_2$, and that (2-23) defines a linear transformation from $\mathcal{V}_1$ to $\mathcal{V}_2$ for *every* choice of these vectors. (Recall that $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is a basis for $\mathcal{V}_1$, and that $x_1, \ldots, x_n$ are the coordinates of $\mathbf{x}$ with respect to this basis.) We now use this observation to define the notion of a matrix for a linear transformation, as follows.

Let $\mathbf{f}_1, \ldots, \mathbf{f}_m$ be a basis for $\mathcal{V}_2$, and let $A\colon \mathcal{V}_1 \to \mathcal{V}_2$ be given. Then, for each integer j, $1 \leq j \leq n$, there exist scalars α_{ij} such that

$$A(\mathbf{e}_j) = \sum_{i=1}^{m} \alpha_{ij}\mathbf{f}_i, \tag{2-24}$$

i.e., such that

$$\begin{aligned} A(\mathbf{e}_1) &= \alpha_{11}\mathbf{f}_1 + \alpha_{21}\mathbf{f}_2 + \cdots + \alpha_{m1}\mathbf{f}_m, \\ A(\mathbf{e}_2) &= \alpha_{12}\mathbf{f}_1 + \alpha_{22}\mathbf{f}_2 + \cdots + \alpha_{m2}\mathbf{f}_m, \\ &\vdots \\ A(\mathbf{e}_n) &= \alpha_{1n}\mathbf{f}_1 + \alpha_{2n}\mathbf{f}_2 + \cdots + \alpha_{mn}\mathbf{f}_m. \end{aligned} \tag{2-25}$$

For computational purposes it turns out to be convenient to display these scalars in the rectangular array

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix}, \tag{2-26}$$

whose *columns* are the coefficients of the various equations in (2–25). It should be noted that the first subscript on an entry in (2–26) indicates the *row* in which that entry appears, and the second indicates the *column*. With this convention in force the entire array can be abbreviated (α_{ij}), it being understood that i and j range independently over the integers $1, \ldots, m$ and $1, \ldots, n$, respectively. When displayed as above, this set of scalars is called *the matrix of A with respect to the bases* $\mathcal{B}_1 = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ *and* $\mathcal{B}_2 = \{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ and is denoted by $[A\colon \mathcal{B}_1, \mathcal{B}_2]$,

or simply by $[A]$. (In the special case where A maps $\mathcal{V}$ into itself and $\mathcal{B}_1 = \mathcal{B}_2 = \mathcal{B}$, the notation $[A\colon \mathcal{B}]$ is also used.) When $m = n$, we say that (2–26) is a *square* matrix; otherwise, *rectangular*. In general, a matrix consisting of m rows and n columns will be referred to as an $m \times n$-matrix (read "m by n").

The argument just given shows that *every* linear transformation from $\mathcal{V}_1$ to $\mathcal{V}_2$ determines a *unique* $m \times n$-matrix with respect to $\mathcal{B}_1$ and $\mathcal{B}_2$. But since the α_{ij} in (2–25) uniquely determine the $A(\mathbf{e}_j)$ and hence, by (2–23), $A(\mathbf{x})$ for all $\mathbf{x}$ in $\mathcal{V}_1$, it follows that *every* $m \times n$-matrix determines a *unique* linear transformation from $\mathcal{V}_1$ to $\mathcal{V}_2$ in terms of $\mathcal{B}_1$ and $\mathcal{B}_2$, and we therefore have

Theorem 2–6. *Let $\mathcal{V}_1$ and $\mathcal{V}_2$ be finite dimensional vector spaces with* $\dim \mathcal{V}_1 = n$, $\dim \mathcal{V}_2 = m$, *and let $\mathcal{B}_1$ and $\mathcal{B}_2$ be bases for $\mathcal{V}_1$ and $\mathcal{V}_2$, respectively. Then every linear transformation from $\mathcal{V}_1$ to $\mathcal{V}_2$ determines a unique $m \times n$-matrix with respect to $\mathcal{B}_1$ and $\mathcal{B}_2$, and, conversely, every such matrix determines a unique linear transformation from $\mathcal{V}_1$ to $\mathcal{V}_2$ defined by* (2–23) *and* (2–24).

It is important to realize that this theorem does *not* assert that every linear transformation has a unique matrix. Indeed, any such assertion would be patently false, for, as we have already seen, the matrix of a linear transformation can change with a change of basis. Thus the several references to bases which appear in Theorem 2–6 cannot under any circumstances be deleted.

EXAMPLE 1. Let $\mathbf{e}_1$ and $\mathbf{e}_2$ be the standard basis vectors in $\mathcal{R}^2$, and let $A: \mathcal{R}^2 \to \mathcal{R}^2$ denote the reflection across the $\mathbf{e}_1$-axis. Then

$$A(\mathbf{e}_1) = 1 \cdot \mathbf{e}_1 + 0 \cdot \mathbf{e}_2, \qquad A(\mathbf{e}_2) = 0 \cdot \mathbf{e}_1 - 1 \cdot \mathbf{e}_2,$$

and the matrix of A with respect to the basis $\mathbf{e}_1, \mathbf{e}_2$ is

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

EXAMPLE 2. Let $A: \mathcal{R}^2 \to \mathcal{R}^2$ be the (counterclockwise) rotation about the origin through an angle θ, and again let $\mathbf{e}_1$ and $\mathbf{e}_2$ be the standard basis vectors. Then

$$\begin{aligned} A(\mathbf{e}_1) &= (\cos\theta)\mathbf{e}_1 + (\sin\theta)\mathbf{e}_2, \\ A(\mathbf{e}_2) &= -(\sin\theta)\mathbf{e}_1 + (\cos\theta)\mathbf{e}_2 \end{aligned}$$

(see Fig. 2–8), and the matrix of A with respect to this basis is

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

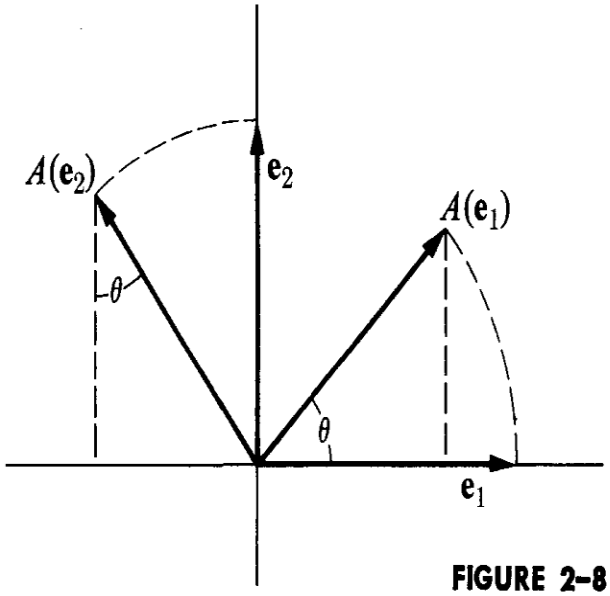

FIGURE 2–8

EXAMPLE 3. If I is the identity map on $\mathcal{V}$, then, regardless of the basis used, the matrix of I is the $n \times n$-matrix

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

with ones along its *principal diagonal* and zeros elsewhere. For obvious reasons this matrix is called the *$n \times n$-identity matrix.*

Similarly the matrix of the zero transformation from $\mathcal{V}_1$ to $\mathcal{V}_2$ is always the the *$m \times n$-zero matrix*

$$\begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}.$$

EXAMPLE 4. Let

$$D\colon \mathcal{P}_n \to \mathcal{P}_n$$

be differentiation, and let

$$\mathcal{B} = \{1, x, \ldots, x^{n-1}\}$$

be the "standard" basis. Then

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

and

$$[D: \mathcal{B}] = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & n-1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

EXAMPLE 5. Let $A: \mathcal{P}_3 \to \mathcal{P}_5$
be the linear transformation defined by

$$A(p(x)) = (2x^2 - 3)p(x)$$

for all $p(x)$ in $\mathcal{P}_3$, and let $\mathcal{B}_1$ and $\mathcal{B}_2$ be the standard bases in $\mathcal{P}_3$ and $\mathcal{P}_5$, respectively. Then

$$A(1) = 2x^2 - 3 = -3 \cdot 1 + 0 \cdot x + 2 \cdot x^2 + 0 \cdot x^3 + 0 \cdot x^4,$$
$$A(x) = 2x^3 - 3x = 0 \cdot 1 - 3 \cdot x + 0 \cdot x^2 + 2 \cdot x^3 + 0 \cdot x^4,$$
$$A(x^2) = 2x^4 - 3x^2 = 0 \cdot 1 + 0 \cdot x - 3 \cdot x^2 + 0 \cdot x^3 + 2 \cdot x^4,$$

and

$$[A: \mathcal{B}_1, \mathcal{B}_2] = \begin{bmatrix} -3 & 0 & 0 \\ 0 & -3 & 0 \\ 2 & 0 & -3 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

ADDITION AND SCALAR MULTIPLICATION OF MATRICES

Let $\mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$ denote the set of all linear transformations from $\mathcal{V}_1$ to $\mathcal{V}_2$, and let $\mathfrak{M}_{mn}$ denote the set of all $m \times n$-matrices. Then if $\dim \mathcal{V}_1 = n$ and $\dim \mathcal{V}_2 = m$, Theorem 2–6 asserts that the function which associates each A in $\mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$ with its matrix $[A]$ with respect to a fixed pair of bases $\mathcal{B}_1$ and $\mathcal{B}_2$ is a *one-to-one* mapping of $\mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$ *onto* $\mathfrak{M}_{mn}$. This simple fact allows us to translate algebraic statements concerning linear transformations into statements concerning matrices, and leads to the subject of matrix algebra. In particular, it allows us to convert $\mathfrak{M}_{mn}$ into a real vector space by using the matrix analogs of the addition and scalar multiplication in $\mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$ to define an addition and scalar multiplication for matrices. The argument goes as follows.

Let (α_{ij}) and (β_{ij}) be arbitrary $m \times n$-matrices, and let σ be a real number. Then by choosing bases in $\mathcal{V}_1$ and $\mathcal{V}_2$ we find *unique* linear transformations A and B in $\mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$ such that

$$[A] = (\alpha_{ij}), \qquad [B] = (\beta_{ij}).$$

Thus, using the notation of the preceding section, we find

$$A(\mathbf{e}_j) = \sum_{i=1}^{m} \alpha_{ij}\mathbf{f}_i, \qquad B(\mathbf{e}_j) = \sum_{i=1}^{m} \beta_{ij}\mathbf{f}_i,$$

and it follows that

$$(A + B)(\mathbf{e}_j) = A(\mathbf{e}_j) + B(\mathbf{e}_j)$$

$$= \sum_{i=1}^{m} \alpha_{ij}\mathbf{f}_i + \sum_{i=1}^{m} \beta_{ij}\mathbf{f}_i = \sum_{i=1}^{m} (\alpha_{ij} + \beta_{ij})\mathbf{f}_i,$$

and

$$(\sigma A)(\mathbf{e}_j) = \sigma\left(\sum_{i=1}^{m} \alpha_{ij}\mathbf{f}_i\right) = \sum_{i=1}^{m} (\sigma\alpha_{ij})\mathbf{f}_i.$$

Hence

$$[A + B] = (\alpha_{ij} + \beta_{ij}),$$

$$[\sigma A] = (\sigma\alpha_{ij}),$$

and if we now require, as reason dictates we must, that

$$[A] + [B] = [A + B], \qquad \text{and} \qquad \sigma[A] = [\sigma A],$$

we are *forced* to give the following definition.

Matrix addition. If *A* and *B* are matrices *of the same size*, then they can be added. (This is similar to the restriction on adding vectors, namely, only vectors from the same space $\mathbf{R}^n$ can be added; you cannot add a 2-vector to a 3-vector, for example.)

Example 1: Consider the following matrices:

$$F = \begin{bmatrix} 2 & -1 \\ 3 & 0 \\ -5 & 2 \end{bmatrix}, \quad G = \begin{bmatrix} 4 & 4 & -3 \\ 0 & -1 & -2 \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 6 \\ -1 & -2 \\ 0 & -3 \end{bmatrix}$$

Which two can be added? What is their sum?

Since only matrices of the same size can be added, only the sum *F* + *H* is defined (*G* cannot be added to either *F* or *H*). The sum of *F* and *H* is

$$F + H = \begin{bmatrix} 2 & -1 \\ 3 & 0 \\ -5 & 2 \end{bmatrix} + \begin{bmatrix} 1 & 6 \\ -1 & -2 \\ 0 & -3 \end{bmatrix} = \begin{bmatrix} 2+1 & -1+6 \\ 3-1 & 0-2 \\ -5+0 & 2-3 \end{bmatrix} = \begin{bmatrix} 3 & 5 \\ 2 & -2 \\ -5 & -1 \end{bmatrix}$$

Since addition of real numbers is commutative, it follows that addition of matrices (when it is defined) is also commutative; that is, for any matrices *A* and *B* of the same size, *A* + *B* will always equal *B* + *A*.

Example 2: If any matrix *A* is added to the zero matrix of the same size, the result is clearly equal to *A*:

A+O=A

This is the matrix analog of the statement *a* + 0 = 0 + *a* = *a*, which expresses the fact that the number 0 is the additive identity in the set of real numbers.

Example 3: Find the matrix *B* such that *A* + *B* = *C*, where

$$A=\begin{bmatrix} 2 & 0 \\ 1 & 4 \end{bmatrix} \quad \text{and} \quad C=\begin{bmatrix} 3 & -1 \\ -2 & 2 \end{bmatrix}$$

if

$$B=\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

then the matrix equation *A* + *B* = *C* becomes

$$\begin{bmatrix} 2+b_{11} & 0+b_{12} \\ 1+b_{21} & 4+b_{22} \end{bmatrix}=\begin{bmatrix} 3 & -1 \\ -2 & 2 \end{bmatrix}$$

Since two matrices are equal if and only if they are of the same size and their corresponding entries are equal, this last equation implies

$$B=\begin{bmatrix} 1 & -1 \\ -3 & -2 \end{bmatrix}$$

This example motivates the definition of matrix *subtraction*.
If *A* and *B* are matrices of the same size, then the entries of *A* − *B* are found by simply subracting the entries of *B* from the corresponding entries of *A*. Since the equation *A* + *B* = *C* is equivalent to *B* = *C* − *A*, employing matrix subtraction above would yield the same result:

$$B = C - A = \begin{bmatrix} 3 & -1 \\ -2 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 3-2 & -1-0 \\ -2-1 & 2-4 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -3 & -2 \end{bmatrix}$$

Definition 2-7. The *sum* $(\alpha_{ij}) + (\beta_{ij})$ of two $m \times n$-matrices is by definition the $m \times n$-matrix $(\alpha_{ij} + \beta_{ij})$; the *product* $\sigma(\alpha_{ij})$ of a real number σ and an $m \times n$-matrix is by definition the $m \times n$-matrix $(\sigma\alpha_{ij})$. In other words,

$$(\alpha_{ij}) + (\beta_{ij}) = (\alpha_{ij} + \beta_{ij}) \tag{2–27}$$

and

$$\sigma(\alpha_{ij}) = (\sigma\alpha_{ij}) \tag{2–28}$$

for all $m \times n$-matrices (α_{ij}), (β_{ij}), and all real numbers σ.

When expressed as rectangular arrays these equations become

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

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

and

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

and assert that matrix addition and scalar multiplication are performed entry by entry, or termwise. Moreover, we now have

Theorem 2–7. *The set $\mathfrak{M}_{mn}$ of all $m \times n$-matrices is a real vector space under the above definition of addition and scalar multiplication.*

The asserted result follows automatically from Theorem 2–6. We can assert that $\mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$ and $\mathfrak{M}_{mn}$ are *algebraically identical* (or *isomorphic*), and that the function which sends each linear transformation $A: \mathcal{V}_1 \to \mathcal{V}_2$ onto its matrix $[A: \mathcal{B}_1, \mathcal{B}_2]$ with respect to a fixed pair of bases $\mathcal{B}_1$ and $\mathcal{B}_2$ is an *isomorphism* of $\mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$ onto $\mathfrak{M}_{mn}$.

As an illustration of the way in which this fact can be used to establish results which are not otherwise obvious, we now propose to show that $\mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$ is finite dimensional and to compute its dimension. For this purpose we introduce the special matrices (e_{ij}), $1 \leq i \leq m$, $1 \leq j \leq n$, each of which has the entry 1 at the intersection of the ith row and jth column and zeros elsewhere:

$$\begin{array}{cc} & \begin{array}{ccc} & j & \end{array} \\ i & \begin{bmatrix} 0 & \vdots & 0 \\ \cdots & 1 & \cdots \\ 0 & \vdots & 0 \end{bmatrix} \end{array} = (e_{ij}). \tag{2–29}$$

Then, for each (α_{ij}) in $\mathfrak{M}_{mn}$ we have

$$\begin{aligned} (\alpha_{ij}) = \quad & \alpha_{11}(e_{11}) + \cdots + \alpha_{1n}(e_{1n}) \\ & + \alpha_{21}(e_{21}) + \cdots + \alpha_{2n}(e_{2n}) \\ & \vdots \\ & + \alpha_{m1}(e_{m1}) + \cdots + \alpha_{mn}(e_{mn}), \end{aligned}$$

or

$$(\alpha_{ij}) = \sum_{i,j=1}^{m,n} \alpha_{ij}(e_{ij}). \tag{2–30}$$

Thus the (e_{ij}) span $\mathfrak{M}_{mn}$, and since it is clear that (2–30) is the only possible way of writing (α_{ij}) as a linear combination of the (e_{ij}), it follows that these matrices are a *basis* for $\mathfrak{M}_{mn}$. (This particular basis is called the *standard basis* for $\mathfrak{M}_{mn}$.) Hence $\mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$ is also finite dimensional with dimension mn, and we have proved

Theorem 2–8. *If $\mathcal{V}_1$ and $\mathcal{V}_2$ are finite dimensional vector spaces, then so is $\mathcal{L}(\mathcal{V}_1, \mathcal{V}_2)$, and*

$$\dim \mathcal{L}(\mathcal{V}_1, \mathcal{V}_2) = (\dim \mathcal{V}_1)(\dim \mathcal{V}_2). \tag{2–31}$$

EXAMPLE 1. The set of all $1 \times n$-matrices $\mathfrak{M}_{1n}$ is an n-dimensional vector space with

$$\begin{aligned} (e_{11}) &= (1, 0, \ldots, 0) \\ (e_{12}) &= (0, 1, \ldots, 0) \\ &\vdots \\ (e_{1n}) &= (0, 0, \ldots, 1) \end{aligned}$$

as a basis. In this case the (e_{ij}) can be identified with the standard basis vectors in $\mathcal{R}^n$, and when this identification is made $\mathfrak{M}_{1n}$ becomes identical with $\mathcal{R}^n$. Thus $\mathcal{L}(\mathcal{V}, \mathcal{R}^1)$ is isomorphic with $\mathcal{R}^n$, whenever $\dim \mathcal{V} = n$, and we conclude that there are exactly as many linear transformations from $\mathcal{V}$ to $\mathcal{R}^1$ as there are vectors in $\mathcal{R}^n$.

EXAMPLE. Let $\mathbf{e}_1$ and $\mathbf{e}_2$ be the standard basis vectors in $\mathcal{R}^2$, and let A be a linear transformation from $\mathcal{R}^2$ to $\mathcal{R}^1$. Then A is completely determined by the pair of real numbers $A(\mathbf{e}_1)$, $A(\mathbf{e}_2)$, and can therefore be represented by the ordered pair $\big(A(\mathbf{e}_1), A(\mathbf{e}_2)\big)$. Since distinct ordered pairs define distinct linear transformations, it follows that there are exactly as many linear transformations from $\mathcal{R}^2$ to $\mathcal{R}^1$ as there are vectors in $\mathcal{R}^2$.

EXAMPLE 2. Let

$$\mathcal{V}_1 = \mathcal{V}_2 = \mathcal{R}^2.$$

Then $\dim \mathcal{L}(\mathcal{R}^2, \mathcal{R}^2) = 4$, and the (e_{ij}) are four in number:

$$(e_{11}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad (e_{12}) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

$$(e_{21}) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad (e_{22}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Moreover, if E_{ij} denotes the linear transformation corresponding to (e_{ij}) with respect to a fixed basis in $\mathcal{R}^2$, then every linear transformation mapping $\mathcal{R}^2$ into itself can be written uniquely in the form

$$\alpha_{11}E_{11} + \alpha_{12}E_{12} + \alpha_{21}E_{21} + \alpha_{22}E_{22}$$

for suitable scalars $\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}$.

EXAMPLE 3. If A is a nonzero linear transformation mapping a finite dimensional vector space $\mathcal{V}$ into itself, then

$$I, A, A^2, \ldots, A^k, \ldots$$

also map $\mathcal{V}$ into itself, and thus belong to $\mathcal{L}(\mathcal{V}, \mathcal{V})$. But by Theorem 2–8 this set is linearly *dependent* in $\mathcal{L}(\mathcal{V}, \mathcal{V})$. Hence there exists a *smallest* positive integer k such that A^k is linearly dependent on $I, A, \ldots, A^{k-1}$, and it is now easy to show that these transformations are a *basis* for the subspace of $\mathcal{L}(\mathcal{V}, \mathcal{V})$ spanned by the powers of A. In particular, we can write A^k in the form

$$A^k = a_{k-1}A^{k-1} + \cdots + a_1 A + a_0 I,$$

or

$$A^k - a_{k-1}A^{k-1} - \cdots - a_1 A - a_0 = O, \tag{2–32}$$

where O is the zero transformation on $\mathcal{V}$, and it follows that A is a root of the polynomial

$$m_A(x) = x^k - a_{k-1}x^{k-1} - \cdots - a_1 x - a_0. \tag{2–33}$$

Since k was chosen as small as possible in this argument there is no polynomial of lower degree having A as a root. For this reason $m_A(x)$ is called the *minimum polynomial* of A. It can be characterized as *the* polynomial of *least degree with leading coefficient* 1 which has A as a root, and is clearly of degree $\leq n^2$ when $\dim \mathcal{V} = n$. Actually, it can be shown that the degree of $m_A(x)$ does not exceed the dimension of $\mathcal{V}$ for any nonzero transformation $A\colon \mathcal{V} \to \mathcal{V}$. The proof, however, is not easy.

MATRIX MULTIPLICATION

Let

$$B: \mathcal{V}_1 \to \mathcal{V}_2, \qquad A: \mathcal{V}_2 \to \mathcal{V}_3$$

be given, let

$$\mathcal{B}_1 = \{\mathbf{e}_1, \ldots, \mathbf{e}_r\}, \qquad \mathcal{B}_2 = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}, \qquad \mathcal{B}_3 = \{\mathbf{g}_1, \ldots, \mathbf{g}_m\}$$

be bases for $\mathcal{V}_1$, $\mathcal{V}_2$, $\mathcal{V}_3$, respectively, and let

$$B(\mathbf{e}_j) = \sum_{k=1}^{n} \beta_{kj}\mathbf{f}_k,$$

$$A(\mathbf{f}_k) = \sum_{i=1}^{m} \alpha_{ik}\mathbf{g}_i.$$

Then we have

$$\begin{aligned} AB(\mathbf{e}_j) &= A\left(\sum_{k=1}^{n} \beta_{kj}\mathbf{f}_k\right) = \sum_{k=1}^{n} \beta_{kj}A(\mathbf{f}_k) \\ &= \sum_{k=1}^{n} \beta_{kj}\left(\sum_{i=1}^{m} \alpha_{ik}\mathbf{g}_i\right) \\ &= \sum_{i=1}^{m}\left(\sum_{k=1}^{n} \alpha_{ik}\beta_{kj}\right)\mathbf{g}_i, \end{aligned}$$

and the matrix of AB with respect to $\mathcal{B}_1, \mathcal{B}_3$ is the $m \times r$-matrix whose ijth entry is $\sum_{k=1}^{n} \alpha_{ik}\beta_{kj}$. But since

$$[A: \mathcal{B}_2, \mathcal{B}_3] = (\alpha_{ik}), \qquad 1 \leq i \leq m, \ 1 \leq k \leq n,$$

and

$$[B: \mathcal{B}_1, \mathcal{B}_2] = (\beta_{kj}), \qquad 1 \leq k \leq n, \ 1 \leq j \leq r,$$

the requirement that

$$[AB: \mathcal{B}_1, \mathcal{B}_3] = [A: \mathcal{B}_2, \mathcal{B}_3][B: \mathcal{B}_1, \mathcal{B}_2]$$

leads to the following definition of matrix multiplication.

Definition 2–8. The *product* $(\alpha_{ik})(\beta_{kj})$ of an $m \times n$-matrix (α_{ik}) and an $n \times r$-matrix (β_{kj}) is by definition the $m \times r$-matrix

$$(\alpha_{ik})(\beta_{kj}) = \left(\sum_{k=1}^{n} \alpha_{ik}\beta_{kj}\right). \tag{2–34}$$

It is important to notice that the product of two matrices is defined *only when the number of columns in the first matrix is equal to the number of rows in the second;* a restriction which is the matrix analog of the fact that the product of two linear transformations is defined only when the image of the first is contained in the domain of the second. When written in greater detail, Eq. (2–34) becomes

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1r} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2r} \\ \vdots & & & \vdots \\ \beta_{n1} & \beta_{n2} & \cdots & \beta_{nr} \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_{11}\beta_{11} + \cdots + \alpha_{1n}\beta_{n1} & \cdots & \alpha_{11}\beta_{1r} + \cdots + \alpha_{1n}\beta_{nr} \\ \alpha_{21}\beta_{11} + \cdots + \alpha_{2n}\beta_{n1} & \cdots & \alpha_{21}\beta_{1r} + \cdots + \alpha_{2n}\beta_{nr} \\ \vdots & & \vdots \\ \alpha_{m1}\beta_{11} + \cdots + \alpha_{mn}\beta_{n1} & \cdots & \alpha_{m1}\beta_{1r} + \cdots + \alpha_{mn}\beta_{nr} \end{bmatrix},$$

and is easily remembered in terms of the kinesthetic relationship between the rows of the first matrix and the columns of the second.

EXAMPLE 1. If

$$(\alpha_{ik}) = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 2 & -3 \end{bmatrix}, \qquad (\beta_{kj}) = \begin{bmatrix} 1 & 3 \\ -2 & 1 \\ 0 & 4 \end{bmatrix},$$

then $(\alpha_{ik})(\beta_{kj})$ is defined, and we have

$$\begin{aligned}(\alpha_{ik})(\beta_{kj}) &= \begin{bmatrix} 2\cdot 1 + (-1)(-2) + 0\cdot 0 & 2\cdot 3 + (-1)\cdot 1 + 0\cdot 4 \\ 1\cdot 1 + 2\cdot(-2) + (-3)\cdot 0 & 1\cdot 3 + 2\cdot 1 + (-3)\cdot 4 \end{bmatrix} \\ &= \begin{bmatrix} 4 & 5 \\ -3 & -7 \end{bmatrix}.\end{aligned}$$

On the other hand, $(\beta_{kj})(\alpha_{ik})$ is not defined since the number of columns in (β_{kj}) is not equal to the number of rows in (α_{ik}).

EXAMPLE 2. Let

$$(\alpha_{ij}) = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}, \qquad (\beta_{ij}) = \begin{bmatrix} -1 & 2 \\ 1 & 0 \end{bmatrix}.$$

Then

$$(\alpha_{ij})(\beta_{ij}) = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -2 & 6 \\ 3 & -2 \end{bmatrix},$$

$$(\beta_{ij})(\alpha_{ij}) = \begin{bmatrix} -1 & 2 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} -5 & 3 \\ 3 & 1 \end{bmatrix},$$

and we see that *the multiplication of square matrices is noncommutative.*