CHAPTER 5

MATRIX TRANSFORMATIONS

In this chapter we will define and study matrix transformation, a special form of linear transformations. The results obtained here have important applications in engineering, physics, and various branches of mathematics.

5.1. The Concept of Matrix Transformations

In this section we will introduce a special class of functions that arises from matrix multiplication. Such functions are important in the study of linear algebra and have significant applications in engineering.

A matrix equation $A\mathbf{x} = \mathbf{b}$ can be regarded as an action of the matrix A on a vector $\mathbf{x}$ by multiplication to produce a new vector $\mathbf{b}$. For example, the equation

$$\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$$

implies that the matrix

$$A = \begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}$$

transforms the vector

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

into the vector $\mathbf{b} = \begin{bmatrix} 5 \\ 8 \end{bmatrix}$. This action represents a function from $\mathbf{x}$ into $A\mathbf{x}$. More generally, we have the following definition.

- **Transformation**

*A **transformation** T from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a rule that assigns to each vector $\mathbf{x}$ in $\mathbb{R}^n$ a vector*

*$T(\mathbf{x})$ in $\mathbb{R}^m$. The set $\mathbb{R}^n$ is called the **domain** of T and the set $\mathbb{R}^m$ is called the **codomain** of T. If the domain and codomain of T is $\mathbb{R}^n$, then T is called an **operator** on $\mathbb{R}^n$.*

In this section we focus on transformations associated with matrix multiplication.

- **Matrix transformation**

  Let A be an $m \times n$ matrix. The transformation T_A that maps $n \times 1$ matrices to $m \times 1$ matrices by the formula

$$T_A(\mathbf{u}) = A\mathbf{u}$$

  *is called a **matrix transformation** associated with the matrix A. If the transformation T_A maps $n \times 1$ matrices to $n \times 1$ matrices, then it is called a **matrix operator**.*

Example: Describe the action of the matrix transformation T_A on the x-axis and y-axis, where $A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$.

Solution: An element of the x-axis is of the form $\mathbf{u} = (x, 0)$. We have

$$T_A(\mathbf{u}) = A\mathbf{u} = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = \begin{bmatrix} x \\ 2x \end{bmatrix} = x \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Thus the x-axis is mapped to all multiples of the vector $(1, 2)$, that is, is mapped to the line $y = 2x$. Similarly, a typical element of the y-axis is $\mathbf{v} = (0, y)$, and we have

$$T_A(\mathbf{v}) = A\mathbf{v} = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} 0 \\ y \end{bmatrix} = \begin{bmatrix} 3y \\ 6y \end{bmatrix} = y \begin{bmatrix} 3 \\ 6 \end{bmatrix},$$

so the y-axis is mapped to the line $y = 2x$ as well.■

Example: Let L be the set of all points satisfying $y = x + 1$. Find $T_A(L)$, where $A = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$.

Solution: An element of the set L is of the form $\mathbf{u} = (x, x + 1)$. We have

$$T_A(L) = A\mathbf{u} = \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} x \\ x+1 \end{bmatrix} = \begin{bmatrix} 4x+3 \\ 8x+6 \end{bmatrix}.$$

Setting $t = 4x + 3$, we see that an element of $T_A(L)$ is of the form $(t, 2t)$. Thus the line $y = x + 1$ is mapped to the line $y = 2x$.■

Matrix transformations are special linear transformations.

- **Linear transformation**

  *A transformation T is **linear** if for all vectors $\mathbf{u}, \mathbf{v}$ in the domain of T, and scalars c, d, we have*

$$T\left(c\mathbf{u}+d\mathbf{v}\right)=cT\left(\mathbf{u}\right)+dT\left(\mathbf{v}\right).$$

  In other words, T is linear if it preserves linear combinations.

Example: Show that the function $T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)=x$ is linear.

Solution: The domain of T is the set of all 2×1 matrices. Let $\begin{bmatrix} x \\ y \end{bmatrix}$ and $\begin{bmatrix} z \\ w \end{bmatrix}$ be in the domain of T, and c, d be any scalars. Since

$$T\left(c\begin{bmatrix} x \\ y \end{bmatrix}+d\begin{bmatrix} z \\ w \end{bmatrix}\right)=T\left(\begin{bmatrix} cx+dz \\ cy+dw \end{bmatrix}\right)=cx+dz=cT\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)+dT\left(\begin{bmatrix} z \\ w \end{bmatrix}\right),$$

we see that T is linear.■

It is easy to verify that every matrix transformation is a linear transformation. Indeed, we have

$$T_A\left(c\mathbf{u}+d\mathbf{v}\right)=A\left(c\mathbf{u}+d\mathbf{v}\right)=A\left(c\mathbf{u}\right)+A\left(d\mathbf{v}\right)=c\left(A\mathbf{u}\right)+d\left(A\mathbf{v}\right)=cT_A\left(\mathbf{u}\right)+dT_A\left(\mathbf{v}\right)$$

for some scalars c and d. However, every linear transformation need not to be a matrix transformation. For example, consider the transformation $D: C^1\left[a,b\right]\to C\left[a,b\right]$ defined by $D\left(f\right)\left(x\right)=f'\left(x\right)$ for all $x\in\left[a,b\right]$. Since

$$D\left(cf+dg\right)\left(x\right)=\left(cf\left(x\right)+dg\left(x\right)\right)'=cf'\left(x\right)+dg'\left(x\right)=cD\left(f\right)\left(x\right)+dD\left(g\right)\left(x\right),$$

D is linear. But, this transformation does not associate any matrix.

The following result and its proof give a condition whether a given linear function is a matrix transformation.

- **Linear transformations and matrix transformations**

  Every linear function from $\mathbb{R}^n$ to $\mathbb{R}^m$ is a matrix transformation.

Proof: Assume that $T:\mathbb{R}^n\to\mathbb{R}^m$ is a linear transformation. We must show that there exists an $m\times n$ matrix A such that $T\left(\mathbf{u}\right)=A\mathbf{u}$ for every vector $\mathbf{u}=\left(x_1,x_2,\ldots,x_n\right)$ in $\mathbb{R}^n$. Let A be the matrix

$$A=\begin{bmatrix} T\left(\mathbf{e}_1\right) & T\left(\mathbf{e}_2\right) & \cdots & T\left(\mathbf{e}_n\right)\end{bmatrix},$$

where $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ are the standard basis vectors for $\mathbb{R}^n$. We then have

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

and hence

$$\begin{aligned} T(\mathbf{u}) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) = x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n) \\ &= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A\mathbf{u}, \end{aligned}$$

the desired result.■

This result and linearity of matrix transformations imply that there is a one-to-one correspondence between $m \times n$ matrices and matrix transformations in the sense that every $m \times n$ matrix A produces exactly one matrix transformation, and every matrix transformation arises from exactly one $m \times n$ matrix, which is called the **standard matrix** for the transformation.

Example: Find the standard matrix A for the linear function $T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 2x_1 + x_2 \\ x_1 - 3x_2 \\ -x_1 + x_2 \end{bmatrix}$

Solution: We have

$$T(\mathbf{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} \quad \text{and} \quad T(\mathbf{e}_2) = T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -3 \\ 1 \end{bmatrix},$$

so the standard matrix is $A = \begin{bmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & -3 \\ -1 & 1 \end{bmatrix}$.■

Suppose that $T_A : \mathbb{R}^n \to \mathbb{R}^k$ and $T_B : \mathbb{R}^k \to \mathbb{R}^m$ are matrix transformations. If $\mathbf{u}$ is a vector in $\mathbb{R}^n$, then T_A maps this vector into a vector $T_A(\mathbf{u})$ in $\mathbb{R}^k$, and T_B maps that vector into the vector $T_B(T_A(\mathbf{u}))$ in $\mathbb{R}^m$. This process creates a transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ that is called the **composition** of T_B with T_A, denoted by $T_B \circ T_A$.

- **Composition and matrix multiplication**

  The composition of matrix transformations corresponds to the matrix product.

Proof: For all vectors $\mathbf{u}$ and for suitably sized matrices A and B, we have $A(B\mathbf{u}) = (AB)\mathbf{u}$. In the matrix transformations terminology, this means that

$$(T_A \circ T_B)(\mathbf{u}) = T_A(T_B(\mathbf{u})) = T_{AB}(\mathbf{u}).$$

Since this equation is true for all $\mathbf{u}$, we conclude that $T_A \circ T_B = T_{AB}$.■

We note that the composition of matrix transformations is not commutative in general.

- **One-to-one matrix transformations**

  *A matrix transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is said to be **one-to-one** if T_A maps distinct vectors in $\mathbb{R}^n$ into distinct vectors in $\mathbb{R}^m$.*

This idea can be expressed in various ways. For example, T_A is one-to-one if the equality $T_A(\mathbf{u}) = T_B(\mathbf{v})$ implies that $\mathbf{u} = \mathbf{v}$.

- **Kernel and range of matrix transformations**

  *If $T_A : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix transformation, then the set of all vectors in $\mathbb{R}^n$ that T_A maps into $\mathbf{0}$ is called the **kernel** of T_A and is denoted by $\ker(T_A)$. The set of all vectors in $\mathbb{R}^m$ that are images under this transformation of at least one vector in $\mathbb{R}^n$ is called the **range** of T_A and is denoted by $R(T_A)$.*

In short we have

$$\ker(T_A) = \text{null space of } A \quad \text{and} \quad R(T_A) = \text{column space of } A$$

Let A be an $n \times n$ matrix, and let $T_A : \mathbb{R}^n \to \mathbb{R}^n$ be a linear operator. We will now investigate relationships between the invertibility of A and properties of T_A. Recall that the following statements are equivalent:

· A is invertible.

· $A\mathbf{u} = \mathbf{v}$ is consistent for every $n \times 1$ matrix $\mathbf{v}$.

· $A\mathbf{u} = \mathbf{v}$ has exactly one solution for every $n \times 1$ matrix $\mathbf{v}$.

Translating these into the corresponding statements about the linear operator T_A, we deduce that the following statements are equivalent:

· A is invertible.

· For every vector $\mathbf{v}$ in $\mathbb{R}^n$, there is some vector $\mathbf{u}$ in $\mathbb{R}^n$ such that $T_A(\mathbf{u}) = \mathbf{v}$. In other words, the range of T_A is $\mathbb{R}^n$.

· For every vector $\mathbf{v}$ in the range of T_A, there is exactly one vector $\mathbf{u}$ in $\mathbb{R}^n$ such that $T_A(\mathbf{u}) = \mathbf{v}$. In other words, T_A is one-to-one.

Moreover, if T_A is one-to-one, then $T_A(\mathbf{u}) = T_A(\mathbf{v}) \Rightarrow \mathbf{u} = \mathbf{v}$, which implies that

$$T_A(\mathbf{u}) - T_A(\mathbf{v}) = \mathbf{0} \Rightarrow T_A(\mathbf{u} - \mathbf{v}) = \mathbf{0} \Rightarrow T_A(\mathbf{0}) = \mathbf{0}.$$

In other words, $\ker(T_A) = \{\mathbf{0}\}$.

In summary, we establish the following results about linear operators on $\mathbb{R}^n$.

- *If A is an $n \times n$ matrix and $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is the corresponding matrix operator, then the following statements are equivalent:*

  (a) *A is invertible.*

  (b) *The kernel of T_A is $\{\mathbf{0}\}$.*

  (c) *The range of T_A is $\mathbb{R}^n$.*

  (d) *T_A is one-to-one.*

If $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is a one-to-one matrix operator, then it follows from above result that A is invertible. The matrix operator $T_{A^{-1}} : \mathbb{R}^n \to \mathbb{R}^n$ that corresponds to A^{-1} is called the **inverse operator** or, simply the **inverse** of T_A. This terminology is appropriate because T_A and $T_{A^{-1}}$ cancel the effect of each other in the sense that if $\mathbf{v}$ is any vector in $\mathbb{R}^n$, then

$$\begin{aligned} T_A(T_{A^{-1}}(\mathbf{v})) &= AA^{-1}\mathbf{v} = I\mathbf{v} = \mathbf{v} \\ T_{A^{-1}}(T_A(\mathbf{v})) &= A^{-1}A\mathbf{v} = I\mathbf{v} = \mathbf{v}, \end{aligned}$$

or equivalently,

$$\begin{aligned} T_A \circ T_{A^{-1}} &= T_{AA^{-1}} = T_I \\ T_{A^{-1}} \circ T_A &= T_{A^{-1}A} = T_I \end{aligned}$$

$T_A : \mathbb{R}^n \to \mathbb{R}^n$ is a one-to-one matrix operator, and if $T_{A^{-1}} : \mathbb{R}^n \to \mathbb{R}^n$ is its inverse, then the standard matrices for these operators are related by the equation $T_{A^{-1}} = T_A^{-1}$.

Example: Show that the operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by the equations

$$\begin{aligned} w_1 &= 2x_1 + x_2 \\ w_2 &= 3x_1 + 4x_2 \end{aligned}$$

is one-to-one, and find $T^{-1}((w_1, w_2))$.

Solution: The matrix form of these equations is

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

so the standard matrix for T is

$$A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}.$$

This matrix is invertible (so T is one-to-one) and the standard matrix for T^{-1} is

$$A^{-1} = \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} \\ -\frac{3}{5} & \frac{2}{5} \end{bmatrix}.$$

Thus

$$T^{-1}((w_1, w_2)) = \begin{bmatrix} \frac{4}{5} & -\frac{1}{5} \\ -\frac{3}{5} & \frac{2}{5} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \frac{4}{5}w_1 - \frac{1}{5}w_2 \\ -\frac{3}{5}w_1 + \frac{2}{5}w_2 \end{bmatrix},$$

from which it follows that

$$T^{-1}((w_1, w_2)) = \left(\frac{4}{5}w_1 - \frac{1}{5}w_2, -\frac{3}{5}w_1 + \frac{2}{5}w_2\right). \blacksquare$$

As a final result in this section, we can add three more equivalent conditions to the list of equivalences for invertibility.

- **Equivalent statements V**

  If A is an $n \times n$ matrix, then the following statements are equivalent:

  (1) *A is invertible.*

  (2) *$Ax = \mathbf{0}$ has only the trivial solution.*

  (3) *The reduced row echelon form of A is I_n.*

  (4) *A is expressible as a product of elementary matrices.*

  (5) *$Ax = b$ is consistent for every $n \times 1$ matrix b.*

  (6) *$Ax = b$ has exactly one solution for every $n \times 1$ matrix b.*

  (7) $\det(A) \neq 0$.

  (8) $N(A) = \{\mathbf{0}\}$.

  (9) *The kernel of T_A is $\{\mathbf{0}\}$.*

  (10) *The range of T_A is $\mathbb{R}^n$.*

  (11) *T_A is one-to-one.*

5.2. Geometric Matrix Transformations

A number of fundamental matrix operators on $\mathbb{R}^2$ or $\mathbb{R}^3$ are used in real-time rendering, the discipline that graphics specialists and game programmers interested in. We give a few examples of such operators in this section.

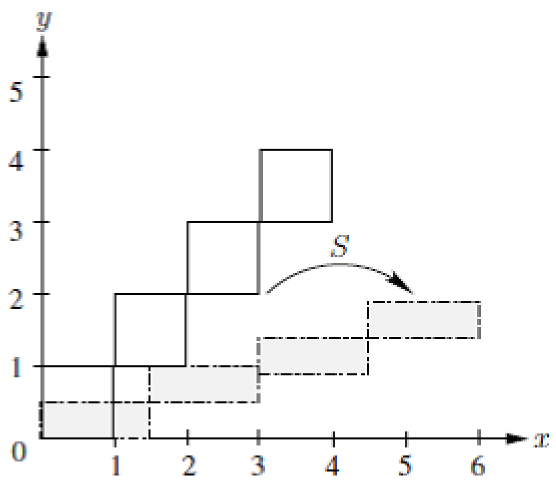
- **Scaling transformation**

  A ***scaling transformation*** *is effected by multiplying each coordinate of a point by a fixed and positive factor.*

Example: The transformation $S((x,y)) = \left(\frac{3}{2}x, \frac{1}{2}y\right)$ is a scaling transformation with

$$S((x,y)) = \begin{bmatrix} \frac{3}{2}x \\ \frac{1}{2}y \end{bmatrix} = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = T_A((x,y)),$$

where $A = \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$.■



- **Contraction and dilation**

  If in a scaling transformation the factor for each coordinate is same, then it is called ***contraction*** *if the factor is less than or equal to 1 and a* ***dilation*** *if the factor is greater than 1.*

- **Shearing transformation**

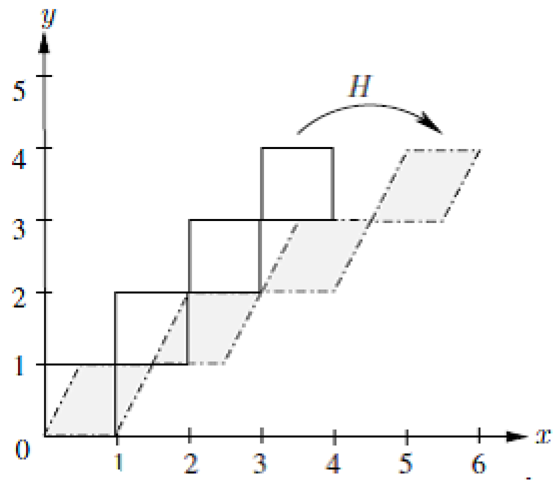
  A ***shearing transformation*** *is effected by adding a constant multiple of one coordinate to another coordinate of the point.*

A shear transformation maps line segments onto line segments and the corners of the square onto the vertices of a parallelogram.

Example: The transformation $H((x,y)) = \left(x + \frac{1}{2}y, y\right)$ is a shearing transformation with

$$H((x,y)) = \begin{bmatrix} x + \frac{1}{2}y \\ y \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = T_B((x,y)),$$

where $B = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{bmatrix}$.■



- **Rotation transformation**

  A ***rotation transformation*** *is effected by rotating each point a fixed angle θ in the counter-clockwise direction about the origin.*

Example: Describe the rotation transformation for the plane.

Solution: We observe that if the point (x,y) is given in polar coordinates as $(r\cos\phi, r\sin\phi)$, then the rotated point (x', y') has coordinates $(r\cos(\phi+\theta), r\sin(\phi+\theta))$. We then have

$$\begin{aligned} \begin{bmatrix} x' \\ y' \end{bmatrix} &= \begin{bmatrix} r\cos(\phi+\theta) \\ r\sin(\phi+\theta) \end{bmatrix} = \begin{bmatrix} r\cos\phi\cos\theta - r\sin\phi\sin\theta \\ r\sin\phi\cos\theta + r\sin\theta\cos\phi \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} r\cos\phi \\ r\sin\phi \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \end{aligned}$$

Defining the rotation matrix $R(\theta)$ by

$$R(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix},$$

we see that $(x', y') = T_{R(\theta)}((x,y))$.■

- **Reflection operators**

  *Linear operators on $\mathbb{R}^2$ or $\mathbb{R}^3$ that map each vector into its symmetric image about some line or plane are called **reflection operators**.*

Operator	Equations	Standard Matrix
Reflection about the y-axis	$\mathbf{v}_1 = -x$ $\mathbf{v}_2 = y$	$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$
Reflection about the y-axis	$\mathbf{v}_1 = x$ $\mathbf{v}_2 = -y$	$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$
Reflection about the line $y = x$	$\mathbf{v}_1 = y$ $\mathbf{v}_2 = x$	$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$
Reflection about the xy-plane	$\mathbf{v}_1 = x$ $\mathbf{v}_2 = y$ $\mathbf{v}_3 = -z$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$
Reflection about the xz-plane	$\mathbf{v}_1 = x$ $\mathbf{v}_2 = -y$ $\mathbf{v}_3 = z$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Reflection about the yz-plane	$\mathbf{v}_1 = -x$ $\mathbf{v}_2 = y$ $\mathbf{v}_3 = z$	$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

- **Projection operator**

  *A **projection operator** (more precisely, an **orthogonal projection operator**) on $\mathbb{R}^2$ or $\mathbb{R}^3$ is any operator that maps each vector into its orthogonal projection on a line or plane through the origin.*

Operator	Equations	Standard Matrix
Orthogonal projection on the x-axis	$\mathbf{v}_1 = x$ $\mathbf{v}_2 = 0$	$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$
Orthogonal projection on the y-axis	$\mathbf{v}_1 = 0$ $\mathbf{v}_2 = y$	$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$
Orthogonal projection on the xy-plane	$\mathbf{v}_1 = x$ $\mathbf{v}_2 = y$ $\mathbf{v}_3 = 0$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
Orthogonal projection on the xz-plane	$\mathbf{v}_1 = x$ $\mathbf{v}_2 = 0$ $\mathbf{v}_3 = z$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Orthogonal projection on the yz-plane	$\mathbf{v}_1 = 0$ $\mathbf{v}_2 = y$ $\mathbf{v}_3 = z$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

5.3. Discrete Dynamical Systems

In many fields such as ecology, economics, and engineering, a need arises to model mathematically a dynamic system that changes over time. Several features of the system are each measured at discrete time intervals, producing a sequence of vectors $x_0, x_1, x_2, \ldots$. The entries in x_k provide information about the state of the system at the time of the kth measurement.

- **Discrete dynamical system**

  *A **discrete linear dynamical system** is a sequence of vectors $\{\mathbf{x}_k\}_{k=0}^{\infty}$, which is defined by an initial vector $\mathbf{x}_0$ and by the rule*

  $$\mathbf{x}_{k+1} = A\mathbf{x}_k,$$

  *where A is a given matrix. The vectors $\mathbf{x}_k$ are called **states**, and the matrix A is called the **transition matrix** of the system.*

Example: Suppose two companies compete for customers in a fixed market in which each customer uses either Brand A or Brand B. Suppose also that a market analysis shows that the buying habits of the customers fit the following pattern in the quarters that were analyzed: each quarter (three-month period), 30% of A users will switch to B, while the rest stay with A. Moreover, 40% of B users will switch to A in a given quarter, while the remaining B users will stay with B. Assuming that this pattern does not vary from quarter to quarter, express the data of this model in matrix terminology.

Solution: If a_0 and b_0 are the fractions of the customers using A and B, respectively, in a given quarter, and a_1 and b_1 the fractions of customers using A and B in the next quarter, then we have

$$\begin{aligned} a_1 &= 0.7a_0 + 0.4b_0 \\ b_1 &= 0.3a_0 + 0.6b_0 \end{aligned}$$

Letting $\mathbf{x}_0 = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix}$ and $\mathbf{x}_1 = \begin{bmatrix} a_1 \\ b_1 \end{bmatrix}$, this system can be written

$$\mathbf{x}_1 = A\mathbf{x}_0,$$

where $A = \begin{bmatrix} 0.7 & 0.4 \\ 0.3 & 0.6 \end{bmatrix}$. Now letting $\mathbf{x}_2 = \begin{bmatrix} a_2 \\ b_2 \end{bmatrix}$, we have

$$\mathbf{x}_2 = A\mathbf{x}_1 = A\left(A\mathbf{x}_0\right) = A^2\mathbf{x}_0.$$

In general, we have

$$\mathbf{x}_k = A^k\mathbf{x}_0,$$

where $\mathbf{x}_k = \begin{bmatrix} a_k \\ b_k \end{bmatrix}$.■

The Markov chains are used as mathematical models of a wide variety of situations in biology, business, chemistry, engineering, physics, and elsewhere. In each case, the model is used to describe an experiment or measurement that is performed many times in the same way, where the outcome of each trial of the experiment will be one of several specified possible outcomes, and where the outcome of one trial depends only on the immediately preceding trial.

- **Markov chain**

  A vector with nonnegative entries that add up to 1 *is called a* ***probability vector****. A* ***stochastic matrix*** *is a square matrix whose columns are probability vectors. A* ***Markov chain*** *is a sequence of probability vectors* $\{\mathbf{x}_k\}$ *with a stochastic matrix* P *such that* $\mathbf{x}_{k+1} = P\mathbf{x}_k$ *for* $k = 0, 1, 2, \ldots$.

Consulting the above example, suppose that Brand A has all the customers initially, that is, the initial state vector is $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Two quarters later we have

$$\begin{bmatrix} a_2 \\ b_2 \end{bmatrix} = \mathbf{x}_2 = A^2\mathbf{x}_0 = \begin{bmatrix} 0.7 & 0.4 \\ 0.3 & 0.6 \end{bmatrix}^2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.61 \\ 0.39 \end{bmatrix}.$$

Thus, Brand A will have 61% of the market and Brand B will have 39% of the market in the second quarter. Similarly, after 20 quarters, we find that

$$\begin{bmatrix} a_{20} \\ b_{20} \end{bmatrix} = \mathbf{x}_{20} = A^{20}\mathbf{x}_0 = \begin{bmatrix} 0.7 & 0.4 \\ 0.3 & 0.6 \end{bmatrix}^{20} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.57143 \\ 0.42857 \end{bmatrix}.$$

Thus, after 20 quarters, Brand A's share will have fallen to about 57% of the market and Brand B's share will have risen to about 43%.

Another important type of model is a so-called **structured population model**. In such a model a population of organisms is divided into a finite number of disjoint states, such as age or weight, so that the entire population is described by a state vector that represents the population at discrete times that occur at a constant period such as every day or year. We give an example in the following.

Example: A certain insect has three life stages: egg, juvenile, and adult. A population is observed in a certain environment to have the following properties in a two-day time span: 20% of the eggs will not survive, and 60% will move to the juvenile stage. In the same time-span 10% of the juveniles will not survive, and 60% will move to the adult stage, while 80% of the adults will survive. Also, in the same time-span adults will product about 0.25 eggs per adult. Assume that initially, there are 10, 8, and 6 eggs, juveniles, and adults (measured in thousands), respectively. Model this population as a discrete dynamical system and use it to compute the population total in 2, 10, and 100 days.

Solution: We start time at day 0 and the kth stage is day $2k$. Here the time period is two days and a state vector has the form $\mathbf{x}_k = \begin{bmatrix} a_k \\ b_k \\ c_k \end{bmatrix}$, where a_k is the number of eggs, b_k the number of juveniles, and c_k the number of adults (all in thousands) on day $2k$. We are given that $\mathbf{x}_0 = \begin{bmatrix} 10 \\ 8 \\ 6 \end{bmatrix}$. Furthermore, the transition matrix has the form

$$A = \begin{bmatrix} 0.2 & 0 & 0.25 \\ 0.6 & 0.3 & 0 \\ 0 & 0.6 & 0.8 \end{bmatrix}.$$

The first column says that 20% of the eggs will remain eggs over one time period, 60% will progress to juveniles, and the rest do not survive. The second column says that juveniles produce no offspring, 30% will remain juveniles, 60% will become adults, and the rest do not survive. The third column says that 0.25 eggs results from one adult, no adult becomes a juvenile, and 80% survive. Now on day 2, we have

$$\mathbf{x}_1 = \begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix} = A\mathbf{x}_0 = \begin{bmatrix} 0.2 & 0 & 0.25 \\ 0.6 & 0.3 & 0 \\ 0 & 0.6 & 0.8 \end{bmatrix} \begin{bmatrix} 10 \\ 8 \\ 6 \end{bmatrix} = \begin{bmatrix} 3.5 \\ 8.4 \\ 9.6 \end{bmatrix},$$

on day 10

$$\mathbf{x}_{10} = \begin{bmatrix} a_{10} \\ b_{10} \\ c_{10} \end{bmatrix} = A^{10}\mathbf{x}_0 = \begin{bmatrix} 0.2 & 0 & 0.25 \\ 0.6 & 0.3 & 0 \\ 0 & 0.6 & 0.8 \end{bmatrix}^{10} \begin{bmatrix} 10 \\ 8 \\ 6 \end{bmatrix} \approx \begin{bmatrix} 3.33 \\ 2.97 \\ 10.3 \end{bmatrix},$$

and on day 100

$$\mathbf{x}_{100} = \begin{bmatrix} a_{100} \\ b_{100} \\ c_{100} \end{bmatrix} = A^{100}\mathbf{x}_0 = \begin{bmatrix} 0.2 & 0 & 0.25 \\ 0.6 & 0.3 & 0 \\ 0 & 0.6 & 0.8 \end{bmatrix}^{100} \begin{bmatrix} 10 \\ 8 \\ 6 \end{bmatrix} \approx \begin{bmatrix} 0.284 \\ 0.253 \\ 0.877 \end{bmatrix}.$$

It appears that the population is declining with time.■