5 Linear Transformations

Definition. If $F: V \to W$ is a function from the vector space V into the vector space W, then F is called a ***linear transformation*** if

(i) $F(\mathbf{u} + \mathbf{v}) = F(\mathbf{u}) + F(\mathbf{v})$ for all vectors $\mathbf{u}$ and $\mathbf{v}$ in V.
(ii) $F(k\mathbf{u}) = kF(\mathbf{u})$ for all vectors $\mathbf{u}$ in V and all scalars k.

To illustrate, let $F: R^2 \to R^3$ be the function defined by (5.1).

$$F(\mathbf{v}) = (x, x + y, x - y) \tag{5.1}$$

$$\begin{aligned} F(\mathbf{u} + \mathbf{v}) &= (x_1 + x_2, [x_1 + x_2] + [y_1 + y_2], [x_1 + x_2] - [y_1 + y_2]) \\ &= (x_1, x_1 + y_1, x_1 - y_1) + (x_2, x_2 + y_2, x_2 - y_2) \\ &= F(\mathbf{u}) + F(\mathbf{v}) \end{aligned}$$

Also, if k is a scalar, $k\mathbf{u} = (kx_1, ky_1)$, so that

$$\begin{aligned} F(k\mathbf{u}) &= (kx_1, kx_1 + ky_1, kx_1 - ky_1) \\ &= k(x_1, x_1 + y_1, x_1 - y_1) \\ &= kF(\mathbf{u}) \end{aligned}$$

Thus F is a linear transformation.

If $F: V \to W$ is a linear transformation, then for any $\mathbf{v}_1$ and $\mathbf{v}_2$ in V and any scalars k_1 and k_2, we have

$$F(k_1\mathbf{v}_1 + k_2\mathbf{v}_2) = F(k_1\mathbf{v}_1) + F(k_2\mathbf{v}_2) = k_1F(\mathbf{v}_1) + k_2F(\mathbf{v}_2)$$

Similarly, if $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ are vectors in V and $k_1, k_2, \ldots, k_n$ are scalars, then

$$F(k_1\mathbf{v}_1 + k_2\mathbf{v}_2 + \cdots + k_n\mathbf{v}_n) = k_1F(\mathbf{v}_1) + k_2F(\mathbf{v}_2) + \cdots + k_nF(\mathbf{v}_n) \tag{5.2}$$

We now give some further examples of linear transformations.

Example 1

Let A be a fixed $m \times n$ matrix. If we use matrix notation for vectors in R^m and R^n, then we can define a function $T: R^n \to R^m$ by

$$T(\mathbf{x}) = A\mathbf{x}$$

Observe that if $\mathbf{x}$ is an $n \times 1$ matrix, then the product $A\mathbf{x}$ is an $m \times 1$ matrix; thus T maps R^n into R^m. Moreover, T is linear; to see this, let $\mathbf{u}$ and $\mathbf{v}$ be $n \times 1$ matrices and let k be a scalar. Using properties of matrix multiplication, we obtain

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} \quad \text{and} \quad A(k\mathbf{u}) = k(A\mathbf{u})$$

or equivalently

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) \quad \text{and} \quad T(k\mathbf{u}) = kT(\mathbf{u})$$

We shall call the linear transformation in this example ***multiplication by A***. Linear transformations of this kind are called ***matrix transformations***.

Example 2

As a special case of the previous example, let θ be a fixed angle, and let $T: R^2 \to R^2$ be multiplication by the matrix

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

If $\mathbf{v}$ is the vector

$$\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$$

then

$$T(\mathbf{v}) = A\mathbf{v} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix}$$

Geometrically, $T(\mathbf{v})$ is the vector that results if $\mathbf{v}$ is rotated through an angle θ. To see this, let ϕ be the angle between $\mathbf{v}$ and the positive x axis, and let

$$\mathbf{v}' = \begin{bmatrix} x' \\ y' \end{bmatrix}$$

be the vector that results when $\mathbf{v}$ is rotated through an angle θ (Figure 5.1). We shall show $\mathbf{v}' = T(\mathbf{v})$. If r denotes the length of $\mathbf{v}$, then

$$x = r\cos\phi \qquad y = r\sin\phi$$

Similarly, since $\mathbf{v}'$ has the same length as $\mathbf{v}$, we have

$$x' = r\cos(\theta + \phi) \qquad y' = r\sin(\theta + \phi)$$

Therefore

$$\begin{aligned} \mathbf{v}' = \begin{bmatrix} x' \\ y' \end{bmatrix} &= \begin{bmatrix} r\cos(\theta + \phi) \\ r\sin(\theta + \phi) \end{bmatrix} \\ &= \begin{bmatrix} r\cos\theta\cos\phi - r\sin\theta\sin\phi \\ r\sin\theta\cos\phi + r\cos\theta\sin\phi \end{bmatrix} \\ &= \begin{bmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} \\ &= A\mathbf{v} = T(\mathbf{v}) \end{aligned}$$

The linear transformation in this example is called the ***rotation of R^2 through the angle θ***.

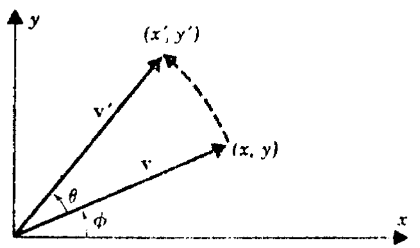

Figure 5.1

5.2 PROPERTIES OF LINEAR TRANSFORMATIONS; KERNEL AND RANGE

Theorem 1. *If $T: V \to W$ is a linear transformation, then:*

(a) $T(\mathbf{0}) = \mathbf{0}$
(b) $T(-\mathbf{v}) = -T(\mathbf{v})$ *for all* $\mathbf{v}$ *in* V
(c) $T(\mathbf{v} - \mathbf{w}) = T(\mathbf{v}) - T(\mathbf{w})$ *for all* $\mathbf{v}$ *and* $\mathbf{w}$ *in* V

Proof. Let $\mathbf{v}$ be any vector in V. Since $0\mathbf{v} = \mathbf{0}$ we have

$$T(\mathbf{0}) = T(0\mathbf{v}) = 0T(\mathbf{v}) = \mathbf{0}$$

which proves (*a*).

Also, $T(-\mathbf{v}) = T((-1)\mathbf{v}) = (-1)T(\mathbf{v}) = -T(\mathbf{v})$, which proves (*b*).

Finally, $\mathbf{v} - \mathbf{w} = \mathbf{v} + (-1)\mathbf{w}$; thus

$$\begin{aligned} T(\mathbf{v} - \mathbf{w}) &= T(\mathbf{v} + (-1)\mathbf{w}) \\ &= T(\mathbf{v}) + (-1)T(\mathbf{w}) \\ &= T(\mathbf{v}) - T(\mathbf{w}) \end{aligned}$$

Definition. If $T: V \to W$ is a linear transformation, then the set of vectors in V that T maps into $\mathbf{0}$ is called the ***kernel*** (or ***nullspace***) of T; it is denoted by $\ker(T)$. The set of all vectors in W that are images under T of at least one vector in V is called the ***range*** of T; it is denoted by $R(T)$.

Example 12

Let $T: V \to W$ be the zero transformation. Since T maps every vector into $\mathbf{0}$, $\ker(T) = V$. Since $\mathbf{0}$ is the only possible image under T, $R(T)$ consists only of the zero vector.

Theorem 2. *If $T: V \to W$ is a linear transformation then:*

(a) The kernel of T is a subspace of V.
(b) The range of T is a subspace of W.

Proof.

(*a*) To show that $\ker(T)$ is a subspace, we must show it is closed under addition and scalar multiplication. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be vectors in $\ker(T)$, and let k be any scalar. Then

$$\begin{aligned} T(\mathbf{v}_1 + \mathbf{v}_2) &= T(\mathbf{v}_1) + T(\mathbf{v}_2) \\ &= \mathbf{0} + \mathbf{0} = \mathbf{0} \end{aligned}$$

so that $\mathbf{v}_1 + \mathbf{v}_2$ is in $\ker(T)$. Also

$$T(k\mathbf{v}_1) = kT(\mathbf{v}_1) = k\mathbf{0} = \mathbf{0}$$

so that $k\mathbf{v}_1$ is in $\ker(T)$.

P. 23c

(*b*) Let $\mathbf{w}_1$ and $\mathbf{w}_2$ be vectors in the range of T. To prove this part we must show that $\mathbf{w}_1 + \mathbf{w}_2$ and $k\mathbf{w}_1$ are in the range of T for any scalar k; that is, we must find vectors $\mathbf{a}$ and $\mathbf{b}$ in V such that $T(\mathbf{a}) = \mathbf{w}_1 + \mathbf{w}_2$ and $T(\mathbf{b}) = k\mathbf{w}_1$.

Since $\mathbf{w}_1$ and $\mathbf{w}_2$ are in the range of T, there are vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ in V such that $T(\mathbf{a}_1) = \mathbf{w}_1$ and $T(\mathbf{a}_2) = \mathbf{w}_2$. Let $\mathbf{a} = \mathbf{a}_1 + \mathbf{a}_2$ and $\mathbf{b} = k\mathbf{a}_1$. Then

$$T(\mathbf{a}) = T(\mathbf{a}_1 + \mathbf{a}_2) = T(\mathbf{a}_1) + T(\mathbf{a}_2) = \mathbf{w}_1 + \mathbf{w}_2$$

and

$$T(\mathbf{b}) = T(k\mathbf{a}_1) = kT(\mathbf{a}_1) = k\mathbf{w}_1$$

which completes the proof.

Example 15

Consider the basis $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for R^3, where $\mathbf{v}_1 = (1, 1, 1)$, $\mathbf{v}_2 = (1, 1, 0)$, $\mathbf{v}_3 = (1, 0, 0)$, and let $T: R^3 \to R^2$ be a linear transformation such that

$$T(\mathbf{v}_1) = (1, 0) \qquad T(\mathbf{v}_2) = (2, -1) \qquad T(\mathbf{v}_3) = (4, 3)$$

Find $T(2, -3, 5)$.

Solution. We first express $\mathbf{v} = (2, -3, 5)$ as a linear combination of $\mathbf{v}_1 = (1, 1, 1)$, $\mathbf{v}_2 = (1, 1, 0)$, and $\mathbf{v}_3 = (1, 0, 0)$. Thus

$$(2, -3, 5) = k_1(1, 1, 1) + k_2(1, 1, 0) + k_3(1, 0, 0)$$

or on equating corresponding components

$$\begin{aligned} k_1 + k_2 + k_3 &= 2 \\ k_1 + k_2 &= -3 \\ k_1 &= 5 \end{aligned}$$

which yields $k_1 = 5$, $k_2 = -8$, $k_3 = 5$ so that

$$(2, -3, 5) = 5\mathbf{v}_1 - 8\mathbf{v}_2 + 5\mathbf{v}_3$$

Thus

$$\begin{aligned} T(2, -3, 5) &= 5T(\mathbf{v}_1) - 8T(\mathbf{v}_2) + 5T(\mathbf{v}_3) \\ &= 5(1, 0) - 8(2, -1) + 5(4, 3) \\ &= (9, 23) \end{aligned}$$

p24a

$$T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 2x_2 \\ x_1 - x_2 \end{bmatrix}$$

then

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

$$A = \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix}$$
$$\begin{matrix} \uparrow & \uparrow \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) \end{matrix}$$

More generally, if

$$T(\mathbf{e}_1) = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}, T(\mathbf{e}_2) = \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}, \ldots, T(\mathbf{e}_n) = \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

then

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \tag{5.6}$$
$$\begin{matrix} \uparrow & \uparrow & & \uparrow \\ T(\mathbf{e}_1) & T(\mathbf{e}_2) & \cdots & T(\mathbf{e}_n) \end{matrix}$$

We shall show that the linear transformation $T: R^n \to R^m$ is multiplication by A. To see this, observe first that

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$$

Therefore, by the linearity of T,

$$T(\mathbf{x}) = x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) + \cdots + x_n T(\mathbf{e}_n) \tag{5.7}$$

P 24b

On the other hand

$$A\mathbf{x} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}$$

$$= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix}$$

$$= x_1 T(\mathbf{e}_1) + x_2 T(\mathbf{e}_2) + \cdots + x_n T(\mathbf{e}_n) \tag{5.8}$$

Comparing (5.7) and (5.8) yields $T(\mathbf{x}) = A\mathbf{x}$, that is, T is multiplication by A.

We shall refer to the matrix A in (5.6) as the ***standard matrix for T***.

Example 19

Find the standard matrix for the transformation $T: R^3 \to R^4$ defined by

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ x_1 - x_2 \\ x_3 \\ x_1 \end{bmatrix}$$

Solution.

$$T(\mathbf{e}_1) = T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \quad T(\mathbf{e}_2) = T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \quad T(\mathbf{e}_3) = T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Using $T(\mathbf{e}_1)$, $T(\mathbf{e}_2)$, and $T(\mathbf{e}_3)$ as column vectors, we obtain

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

As a check, observe that

$$A\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_1 - x_2 \\ x_3 \\ x_1 \end{bmatrix}$$

which agrees with the given formula for T.

6 Eigenvalues, Eigenvectors

Definition. If A is an $n \times n$ matrix, then a nonzero vector $\mathbf{x}$ in R^n is called an ***eigenvector*** of A if $A\mathbf{x}$ is a scalar multiple of $\mathbf{x}$; that is

$$A\mathbf{x} = \lambda \mathbf{x}$$

for some scalar λ. The scalar λ is called an ***eigenvalue*** of A and $\mathbf{x}$ is said to be an eigenvector ***corresponding*** to λ.

To find the eigenvalues of an $n \times n$ matrix A we rewrite $A\mathbf{x} = \lambda\mathbf{x}$ as

$$A\mathbf{x} = \lambda I\mathbf{x}$$

or equivalently

$$(\lambda I - A)\mathbf{x} = 0 \qquad (6.1)$$

For λ to be an eigenvalue, there must be a nonzero solution of this equation. However, by Theorem 13 of Section 4.6, Equation 6.1 will have a nonzero solution if and only if

$$\det(\lambda I - A) = 0$$

This is called the ***characteristic equation*** of A; the scalars satisfying this equation are the eigenvalues of A.

Example 1

The vector $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is an eigenvector of

$$A = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}$$

corresponding to the eigenvalue $\lambda = 3$ since

$$A\mathbf{x} = \begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix} = 3\mathbf{x}$$

Example 2

Find the eigenvalues of the matrix

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

Solution. Since

$$\lambda I - A = \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 3 & 2 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} \lambda - 3 & -2 \\ 1 & \lambda \end{bmatrix}$$

and

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 3 & -2 \\ 1 & \lambda \end{bmatrix} = \lambda^2 - 3\lambda + 2$$

the characteristic equation of A is

$$\lambda^2 - 3\lambda + 2 = 0$$

The solutions of this equation are $\lambda = 1$ and $\lambda = 2$; these are the eigenvalues of A.

Example 4

Find the eigenvalues of

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

Solution. As in the preceding examples

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda - 2 & -1 & 0 \\ -3 & \lambda - 2 & 0 \\ 0 & 0 & \lambda - 4 \end{bmatrix} = \lambda^3 - 8\lambda^2 + 17\lambda - 4$$

The eigenvalues of A must therefore satisfy the cubic equation

$$\lambda^3 - 8\lambda^2 + 17\lambda - 4 = 0 \qquad (6.2)$$

$$(\lambda - 4)(\lambda^2 - 4\lambda + 1) = 0$$

$$\lambda = 4 \qquad \lambda = 2 + \sqrt{3} \qquad \text{and} \qquad \lambda = 2 - \sqrt{3}$$

P25a

Theorem 1. *If A is an $n \times n$ matrix, then the following are equivalent.*

(*a*) *λ is an eigenvalue of A.*
(*b*) *The system of equations $(\lambda I - A)\mathbf{x} = \mathbf{0}$ has nontrivial solutions.*
(*c*) *There is a nonzero vector $\mathbf{x}$ in R^n such that $A\mathbf{x} = \lambda\mathbf{x}$.*
(*d*) *λ is a real solution of the characteristic equation $\det(\lambda I - A) = 0$.*

Now that we know how to find eigenvalues we turn to the problem of finding eigenvectors. The eigenvectors of A corresponding to an eigenvalue λ are the nonzero vectors that satisfy $A\mathbf{x} = \lambda\mathbf{x}$. Equivalently the eigenvectors corresponding to λ are the nonzero vectors in the solution space of $(\lambda I - A)\mathbf{x} = \mathbf{0}$. We call this solution space the ***eigenspace*** of A corresponding to λ.

Example 5

Find bases for the eigenspaces of

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

Solution. The characteristic equation of A is $(\lambda - 1)(\lambda - 5)^2 = 0$ (verify), so that the eigenvalues of A are $\lambda = 1$ and $\lambda = 5$.

By definition

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

is an eigenvector of A corresponding to λ if and only if $\mathbf{x}$ is a nontrivial solution of $(\lambda I - A)\mathbf{x} = \mathbf{0}$, that is, of

$$\begin{bmatrix} \lambda - 3 & 2 & 0 \\ 2 & \lambda - 3 & 0 \\ 0 & 0 & \lambda - 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{6.3}$$

P25b

If $\lambda = 5$, (6.3) becomes

$$\begin{bmatrix} 2 & 2 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving this system yields (verify)

$$x_1 = -s \qquad x_2 = s \qquad x_3 = t$$

Thus the eigenvectors of A corresponding to $\lambda = 5$ are the nonzero vectors of the form

$$\mathbf{x} = \begin{bmatrix} -s \\ s \\ t \end{bmatrix} = \begin{bmatrix} -s \\ s \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ t \end{bmatrix} = s\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Since

$$\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}. \quad \text{and} \quad \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

are linearly independent, they form a basis for the eigenspace corresponding to $\lambda = 5$.

If $\lambda = 1$, then (6.3) becomes

$$\begin{bmatrix} -2 & 2 & 0 \\ 2 & -2 & 0 \\ 0 & 0 & -4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Solving this system yields (verify)

$$x_1 = t \qquad x_2 = t \qquad x_3 = 0$$

Thus the eigenvectors corresponding to $\lambda = 1$ are the nonzero vectors of the form

$$\mathbf{x} = \begin{bmatrix} t \\ t \\ 0 \end{bmatrix} = t\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

so that

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

is a basis for the eigenspace corresponding to $\lambda = 1$.

P25C

P25d

Example 7

Find a matrix P that diagonalizes

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

Solution. From Example 5 the eigenvalues of A are $\lambda = 1$ and $\lambda = 5$. Also from that example the vectors

$$\mathbf{p}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{p}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

form a basis for the eigenspace corresponding to $\lambda = 5$ and

$$\mathbf{p}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

is a basis for the eigenspace corresponding to $\lambda = 1$. It is easy to check that $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ is linearly independent, so that

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

diagonalizes A. As a check, the reader should verify that

$$P^{-1}AP = \begin{bmatrix} -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

There is no preferred order for the columns of P. Since the ith diagonal entry of $P^{-1}AP$ is an eigenvalue for the ith column vector of P, changing the order of the columns of P just changes the order of the eigenvalues on the diagonal of $P^{-1}AP$. Thus, had we written

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

in the last example, we would have obtained

$$P^{-1}AP = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

Example 8

The characteristic equation of

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

is

$$\det(\lambda I - A) = \det \begin{bmatrix} \lambda + 3 & -2 \\ 2 & \lambda - 1 \end{bmatrix} = (\lambda + 1)^2 = 0$$

Thus $\lambda = -1$ is the only eigenvalue of A; the eigenvectors corresponding to $\lambda = -1$ are the solutions of $(-I - A)\mathbf{x} = \mathbf{0}$; that is, of

$$\begin{aligned} 2x_1 - 2x_2 &= 0 \\ 2x_1 - 2x_2 &= 0 \end{aligned}$$

The solutions of this system are $x_1 = t, x_2 = t$ (verify); hence the eigenspace consists of all vectors of the form

$$\begin{bmatrix} t \\ t \end{bmatrix} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Since this space is 1-dimensional, A does not have two linearly independent eigenvectors, and is therefore not diagonalizable.