Chapter 3

Linear Algebra

3.1 Introduction

This chapter reviews a number of concepts and results in linear algebra that are essential in the study of this text. The topics are carefully selected, and only those that will be used subsequently are introduced. Most results are developed intuitively in order for the reader to better grasp the ideas. They are stated as theorems for easy reference in later chapters. However, no formal proofs are given.

As we saw in the preceding chapter, all parameters that arise in the real world are real numbers. Therefore we deal only with real numbers, unless stated otherwise, throughout this text. Let **A**, **B**, **C**, and **D** be, respectively, $n \times m$, $m \times r$, $l \times n$, and $r \times p$ real matrices. Let $\mathbf{a}_i$ be the ith column of **A**, and $\mathbf{b}_j$ the jth row of **B**. Then we have

$$\mathbf{AB} = [\mathbf{a}_1 \;\; \mathbf{a}_2 \;\; \cdots \;\; \mathbf{a}_m] \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_m \end{bmatrix} = \mathbf{a}_1\mathbf{b}_1 + \mathbf{a}_2\mathbf{b}_2 + \cdots + \mathbf{a}_m\mathbf{b}_m \tag{3.1}$$

$$\mathbf{CA} = \mathbf{C}\,[\mathbf{a}_1 \;\; \mathbf{a}_2 \;\; \cdots \;\; \mathbf{a}_m] = [\mathbf{Ca}_1 \;\; \mathbf{Ca}_2 \;\; \cdots \;\; \mathbf{Ca}_m] \tag{3.2}$$

and

$$\mathbf{BD} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_m \end{bmatrix} \mathbf{D} = \begin{bmatrix} \mathbf{b}_1\mathbf{D} \\ \mathbf{b}_2\mathbf{D} \\ \vdots \\ \mathbf{b}_m\mathbf{D} \end{bmatrix} \tag{3.3}$$

These identities can easily be verified. Note that $\mathbf{a}_i\mathbf{b}_i$ is an $n \times r$ matrix; it is the product of an $n \times 1$ column vector and a $1 \times r$ row vector. The product $\mathbf{b}_i\mathbf{a}_i$ is not defined unless $n = r$; it becomes a scalar if $n = r$.

3.2 Basis, Representation, and Orthonormalization

Consider an n-dimensional real linear space, denoted by $\mathcal{R}^n$. Every vector in $\mathcal{R}^n$ is an n-tuple of real numbers such as

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

To save space, we write it as $\mathbf{x} = [x_1 \;\; x_2 \;\; \cdots \;\; x_n]'$, where the prime denotes the transpose.

The set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ in $\mathcal{R}^n$ is said to be *linearly dependent* if there exist real numbers $\alpha_1, \alpha_2, \ldots, \alpha_m$, not all zero, such that

$$\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_m\mathbf{x}_m = \mathbf{0} \tag{3.4}$$

If the only set of α_i for which (3.4) holds is $\alpha_1 = 0, \; \alpha_2 = 0, \; \ldots, \; \alpha_m = 0$, then the set of vectors $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}$ is said to be *linearly independent*.

If the set of vectors in (3.4) is linearly dependent, then there exists at least one α_i, say, α_1, that is different from zero. Then (3.4) implies

$$\begin{aligned} \mathbf{x}_1 &= -\frac{1}{\alpha_1}[\alpha_2\mathbf{x}_2 + \alpha_3\mathbf{x}_3 + \cdots + \alpha_m\mathbf{x}_m] \\ &=: \beta_2\mathbf{x}_2 + \beta_3\mathbf{x}_3 + \cdots + \beta_m\mathbf{x}_m \end{aligned}$$

where $\beta_i = -\alpha_i/\alpha_1$. Such an expression is called a linear combination.

The *dimension* of a linear space can be defined as the maximum number of linearly independent vectors in the space. Thus in $\mathcal{R}^n$, we can find at most n linearly independent vectors.

Basis and representation A set of linearly independent vectors in $\mathcal{R}^n$ is called a *basis* if every vector in $\mathcal{R}^n$ can be expressed as a unique linear combination of the set. In $\mathcal{R}^n$, any set of n linearly independent vectors can be used as a basis. Let $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ be such a set. Then every vector $\mathbf{x}$ can be expressed uniquely as

$$\mathbf{x} = \alpha_1\mathbf{q}_1 + \alpha_2\mathbf{q}_2 + \cdots + \alpha_n\mathbf{q}_n \tag{3.5}$$

Define the $n \times n$ square matrix

$$\mathbf{Q} := [\mathbf{q}_1 \;\; \mathbf{q}_2 \;\; \cdots \;\; \mathbf{q}_n] \tag{3.6}$$

Then (3.5) can be written as

$$\mathbf{x} = \mathbf{Q}\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} =: \mathbf{Q}\bar{\mathbf{x}} \tag{3.7}$$

We call $\bar{\mathbf{x}} = [\alpha_1 \ \alpha_2 \ \cdots \ \alpha_n]'$ the *representation* of the vector $\mathbf{x}$ with respect to the basis $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$.

We will associate with every $\mathcal{R}^n$ the following *orthonormal basis*:

$$\mathbf{i}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{i}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad \cdots, \quad \mathbf{i}_{n-1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{i}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{3.8}$$

With respect to this basis, we have

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

where $\mathbf{I}_n$ is the $n \times n$ unit matrix. In other words, the representation of any vector $\mathbf{x}$ with respect to the orthonormal basis in (3.8) equals itself.

EXAMPLE 3.1 Consider the vector $\mathbf{x} = [1 \ \ 3]'$ in $\mathcal{R}^2$ as shown in Fig. 3.1, The two vectors $\mathbf{q}_1 = [3 \ \ 1]'$ and $\mathbf{q}_2 = [2 \ \ 2]'$ are clearly linearly independent and can be used as a basis. If we draw from $\mathbf{x}$ two lines in parallel with $\mathbf{q}_2$ and $\mathbf{q}_1$, they intersect at $-\mathbf{q}_1$ and $2\mathbf{q}_2$ as shown. Thus the representation of $\mathbf{x}$ with respect to $\{\mathbf{q}_1, \mathbf{q}_2\}$ is $[-1 \ \ 2]'$. This can also be verified from

$$\mathbf{x} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} = [\mathbf{q}_1 \ \ \mathbf{q}_2]\begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

To find the representation of $\mathbf{x}$ with respect to the basis $\{\mathbf{q}_2, \mathbf{i}_2\}$, we draw from $\mathbf{x}$ two lines in parallel with $\mathbf{i}_2$ and $\mathbf{q}_2$. They intersect at $0.5\mathbf{q}_2$ and $2\mathbf{i}_2$. Thus the representation of $\mathbf{x}$ with respect to $\{\mathbf{q}_2, \mathbf{i}_2\}$ is $[0.5 \ \ 2]'$. (Verify.)

Norms of vectors The concept of *norm* is a generalization of length or magnitude. Any real-valued function of $\mathbf{x}$, denoted by $||\mathbf{x}||$, can be defined as a norm if it has the following properties:

1. $||\mathbf{x}|| \geq 0$ for every $\mathbf{x}$ and $||\mathbf{x}|| = 0$ if and only if $\mathbf{x} = \mathbf{0}$.
2. $||\alpha\mathbf{x}|| = |\alpha|||\mathbf{x}||$, for any real α.
3. $||\mathbf{x}_1 + \mathbf{x}_2|| \leq ||\mathbf{x}_1|| + ||\mathbf{x}_2||$ for every $\mathbf{x}_1$ and $\mathbf{x}_2$.

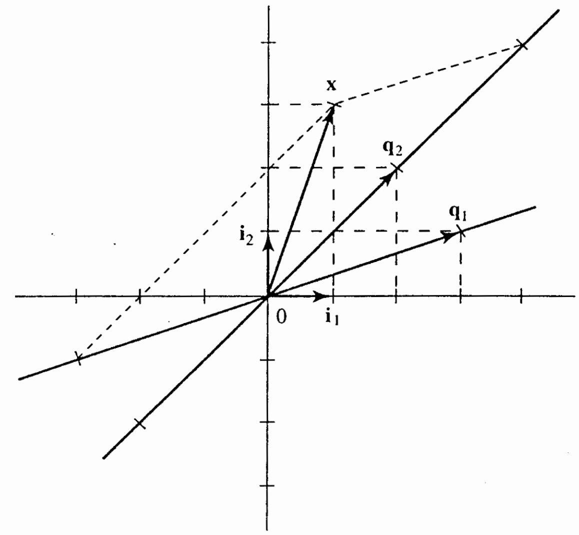

Figure 3.1 Different representations of vector **x**.

The last inequality is called the *triangular inequality*.

Let $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_n]'$. Then the norm of **x** can be chosen as any one of the following:

$$||\mathbf{x}||_1 := \sum_{i=1}^{n} |x_i|$$

$$||\mathbf{x}||_2 := \sqrt{\mathbf{x}'\mathbf{x}} = \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2}$$

$$||\mathbf{x}||_\infty := \max_i \ |x_i|$$

They are called, respectively, 1-norm, 2- or Euclidean norm, and infinite-norm. The 2-norm is the length of the vector from the origin. We use exclusively, unless stated otherwise, the Euclidean norm and the subscript 2 will be dropped.

In MATLAB, the norms just introduced can be obtained by using the functions `norm(x,1)`, `norm(x,2)=norm(x)`, and `norm(x,inf)`.

Orthonormalization A vector **x** is said to be normalized if its Euclidean norm is 1 or $\mathbf{x}'\mathbf{x} = 1$. Note that $\mathbf{x}'\mathbf{x}$ is scalar and $\mathbf{x}\mathbf{x}'$ is $n \times n$. Two vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ are said to be *orthogonal* if $\mathbf{x}_1'\mathbf{x}_2 = \mathbf{x}_2'\mathbf{x}_1 = 0$. A set of vectors $\mathbf{x}_i$, $i = 1, 2, \ldots, m$, is said to be *orthonormal* if

$$\mathbf{x}_i'\mathbf{x}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

Given a set of linearly independent vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_m$, we can obtain an orthonormal set using the procedure that follows:

$$\mathbf{u}_1 := \mathbf{e}_1 \qquad\qquad \mathbf{q}_1 := \mathbf{u}_1/||\mathbf{u}_1||$$
$$\mathbf{u}_2 := \mathbf{e}_2 - (\mathbf{q}_1'\mathbf{e}_2)\mathbf{q}_1 \qquad\qquad \mathbf{q}_2 := \mathbf{u}_2/||\mathbf{u}_2||$$
$$\vdots$$
$$\mathbf{u}_m := \mathbf{e}_m - \textstyle\sum_{k=1}^{m-1}(\mathbf{q}_k'\mathbf{e}_m)\mathbf{q}_k \qquad\qquad \mathbf{q}_m := \mathbf{u}_m/||\mathbf{u}_m||$$

The first equation normalizes the vector $\mathbf{e}_1$ to have norm 1. The vector $(\mathbf{q}_1'\mathbf{e}_2)\mathbf{q}_1$ is the projection of the vector $\mathbf{e}_2$ along $\mathbf{q}_1$. Its subtraction from $\mathbf{e}_2$ yields the vertical part $\mathbf{u}_2$. It is then normalized to 1 as shown in Fig. 3.2. Using this procedure, we can obtain an orthonormal set. This is called the *Schmidt orthonormalization procedure*.

Let $\mathbf{A} = [\mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_m]$ be an $n \times m$ matrix with $m \leq n$. If all columns of $\mathbf{A}$ or $\{\mathbf{a}_i,\ i = 1, 2, \ldots, m\}$ are orthonormal, then

$$\mathbf{A}'\mathbf{A} = \begin{bmatrix} \mathbf{a}_1' \\ \mathbf{a}_2' \\ \vdots \\ \mathbf{a}_m' \end{bmatrix} [\mathbf{a}_1\ \mathbf{a}_2\ \cdots\ \mathbf{a}_m] = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \mathbf{I}_m$$

where $\mathbf{I}_m$ is the unit matrix of order m. Note that, in general, $\mathbf{A}\mathbf{A}' \neq \mathbf{I}_n$. See Problem 3.4.

3.3 Linear Algebraic Equations

Consider the set of linear algebraic equations

$$\mathbf{A}\mathbf{x} = \mathbf{y} \tag{3.9}$$

where $\mathbf{A}$ and $\mathbf{y}$ are, respectively, $m \times n$ and $m \times 1$ real matrices and $\mathbf{x}$ is an $n \times 1$ vector. The matrices $\mathbf{A}$ and $\mathbf{y}$ are given and $\mathbf{x}$ is the unknown to be solved. Thus the set actually consists of m equations and n unknowns. The number of equations can be larger than, equal to, or smaller than the number of unknowns.

We discuss the existence condition and general form of solutions of (3.9). The *range space* of $\mathbf{A}$ is defined as all possible linear combinations of all columns of $\mathbf{A}$. The *rank* of $\mathbf{A}$ is

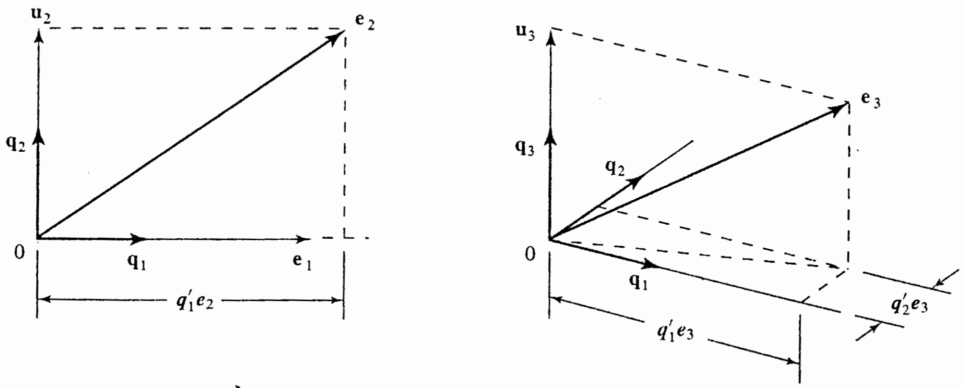

Figure 3.2 Schmidt orthonormization procedure.

defined as the dimension of the range space or, equivalently, the number of linearly independent columns in **A**. A vector **x** is called a *null vector* of **A** if $\mathbf{Ax} = \mathbf{0}$. The *null space* of **A** consists of all its null vectors. The *nullity* is defined as the maximum number of linearly independent null vectors of **A** and is related to the rank by

$$\text{Nullity }(\mathbf{A}) = \text{ number of columns of } \mathbf{A} - \text{ rank }(\mathbf{A}) \tag{3.10}$$

Example 3.2 Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 2 & 3 & 4 \\ 2 & 0 & 2 & 0 \end{bmatrix} =: [\mathbf{a}_1 \;\; \mathbf{a}_2 \;\; \mathbf{a}_3 \;\; \mathbf{a}_4] \tag{3.11}$$

where $\mathbf{a}_i$ denotes the ith column of **A**. Clearly $\mathbf{a}_1$ and $\mathbf{a}_2$ are linearly independent. The third column is the sum of the first two columns or $\mathbf{a}_1 + \mathbf{a}_2 - \mathbf{a}_3 = \mathbf{0}$. The last column is twice the second column, or $2\mathbf{a}_2 - \mathbf{a}_4 = \mathbf{0}$. Thus **A** has two linearly independent columns and has rank 2. The set $\{\mathbf{a}_1, \; \mathbf{a}_2\}$ can be used as a basis of the range space of **A**.

Equation (3.10) implies that the nullity of **A** is 2. It can readily be verified that the two vectors

$$\mathbf{n}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix} \qquad \mathbf{n}_2 = \begin{bmatrix} 0 \\ 2 \\ 0 \\ -1 \end{bmatrix} \tag{3.12}$$

meet the condition $\mathbf{An}_i = \mathbf{0}$. Because the two vectors are linearly independent, they form a basis of the null space.

The rank of **A** is defined as the number of linearly independent columns. It also equals the number of linearly independent rows. Because of this fact, if **A** is $m \times n$, then

$$\text{rank}(\mathbf{A}) \le \min(m, \; n)$$

In MATLAB, the range space, null space, and rank can be obtained by calling the functions `orth`, `null`, and `rank`. For example, for the matrix in (3.11), we type

```
a=[0  1 1 2;1 2 3 4;2 0 2 0];
rank(a)
```

which yields 2. Note that MATLAB computes ranks by using singular-value decomposition (svd), which will be introduced later. The svd algorithm also yields the range and null spaces of the matrix. The MATLAB function `R=orth(a)` yields[1]

```
Ans     R=
```

$$\begin{matrix} 0.3782 & -0.3084 \\ 0.8877 & -0.1468 \\ 0.2627 & 0.9399 \end{matrix} \tag{3.13}$$

1. This is obtained using MATLAB Version 5. Earlier versions may yield different results.

The two columns of R form an orthonormal basis of the range space. To check the orthonormality, we type `R'*R`, which yields the unity matrix of order 2. The two columns in R are not obtained from the basis $\{\mathbf{a}_1, \mathbf{a}_2\}$ in (3.11) by using the Schmidt orthonormalization procedure; they are a by-product of svd. However, the two bases should span the same range space. This can be verified by typing

```
rank([a1 a2 R])
```

which yields 2. This confirms that $\{\mathbf{a}_1, \mathbf{a}_2\}$ span the same space as the two vectors of R. We mention that the rank of a matrix can be very sensitive to roundoff errors and imprecise data. For example, if we use the five-digit display of R in (3.13), the rank of `[a1 a2 R]` is 3. The rank is 2 if we use the R stored in MATLAB, which uses 16 digits plus exponent.

The null space of (3.11) can be obtained by typing `null(a)`, which yields

```
Ans    N=
```

$$\begin{matrix} 0.3434 & -0.5802 \\ 0.8384 & 0.3395 \\ -0.3434 & 0.5802 \\ -0.2475 & -0.4598 \end{matrix} \tag{3.14}$$

The two columns are an orthonormal basis of the null space spanned by the two vectors $\{\mathbf{n}_1, \mathbf{n}_2\}$ in (3.12). All discussion for the range space applies here. That is, `rank([n1 n2 N])` yields 3 if we use the five-digit display in (3.14). The rank is 2 if we use the N stored in MATLAB.

With this background, we are ready to discuss solutions of (3.9). We use ρ to denote the rank of a matrix.

Theorem 3.1

1. Given an $m \times n$ matrix $\mathbf{A}$ and an $m \times 1$ vector $\mathbf{y}$, an $n \times 1$ solution $\mathbf{x}$ exists in $\mathbf{Ax} = \mathbf{y}$ if and only if $\mathbf{y}$ lies in the range space of $\mathbf{A}$ or, equivalently,

$$\rho(\mathbf{A}) = \rho([\mathbf{A}\ \ \mathbf{y}])$$

where $[\mathbf{A}\ \ \mathbf{y}]$ is an $m \times (n+1)$ matrix with $\mathbf{y}$ appended to $\mathbf{A}$ as an additional column.

2. Given $\mathbf{A}$, a solution $\mathbf{x}$ exists in $\mathbf{Ax} = \mathbf{y}$ for every $\mathbf{y}$, if and only if $\mathbf{A}$ has rank m (full row rank).

The first statement follows directly from the definition of the range space. If $\mathbf{A}$ has full row rank, then the rank condition in (1) is always satisfied for every $\mathbf{y}$. This establishes the second statement.

Theorem 3.2 (Parameterization of all solutions)

Given an $m \times n$ matrix $\mathbf{A}$ and an $m \times 1$ vector $\mathbf{y}$, let $\mathbf{x}_p$ be a solution of $\mathbf{Ax} = \mathbf{y}$ and let $k := n - \rho(\mathbf{A})$ be the nullity of $\mathbf{A}$. If $\mathbf{A}$ has rank n (full column rank) or $k = 0$, then the solution $\mathbf{x}_p$ is unique. If $k > 0$, then for every real α_i, $i = 1, 2, \ldots, k$, the vector

$$\mathbf{x} = \mathbf{x}_p + \alpha_1 \mathbf{n}_1 + \cdots + \alpha_k \mathbf{n}_k \tag{3.15}$$

is a solution of $\mathbf{Ax} = \mathbf{y}$, where $\{\mathbf{n}_1, \ldots, \mathbf{n}_k\}$ is a basis of the null space of $\mathbf{A}$.

Substituting (3.15) into $\mathbf{Ax} = \mathbf{y}$ yields

$$\mathbf{A}\mathbf{x}_p + \sum_{i=1}^{k} \alpha_i \mathbf{A}\mathbf{n}_i = \mathbf{A}\mathbf{x}_p + \mathbf{0} = \mathbf{y}$$

Thus, for every α_i, (3.15) is a solution. Let $\bar{\mathbf{x}}$ be a solution or $\mathbf{A}\bar{\mathbf{x}} = \mathbf{y}$. Subtracting this from $\mathbf{A}\mathbf{x}_p = \mathbf{y}$ yields

$$\mathbf{A}(\bar{\mathbf{x}} - \mathbf{x}_p) = \mathbf{0}$$

which implies that $\bar{\mathbf{x}} - \mathbf{x}_p$ is in the null space. Thus $\bar{\mathbf{x}}$ can be expressed as in (3.15). This establishes Theorem 3.2.

Example 3.3 Consider the equation

$$\mathbf{Ax} = \begin{bmatrix} 0 & 1 & 1 & 2 \\ 1 & 2 & 3 & 4 \\ 2 & 0 & 2 & 0 \end{bmatrix} \mathbf{x} =: [\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3 \ \mathbf{a}_4]\mathbf{x} = \begin{bmatrix} -4 \\ -8 \\ 0 \end{bmatrix} = \mathbf{y} \tag{3.16}$$

This $\mathbf{y}$ clearly lies in the range space of $\mathbf{A}$ and $\mathbf{x}_p = [0 \ -4 \ 0 \ 0]'$ is a solution. A basis of the null space of $\mathbf{A}$ was shown in (3.12). Thus the general solution of (3.16) can be expressed as

$$\mathbf{x} = \mathbf{x}_p + \alpha_1 \mathbf{n}_1 + \alpha_2 \mathbf{n}_2 = \begin{bmatrix} 0 \\ -4 \\ 0 \\ 0 \end{bmatrix} + \alpha_1 \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 2 \\ 0 \\ -1 \end{bmatrix} \tag{3.17}$$

for any real α_1 and α_2.

In application, we will also encounter $\mathbf{xA} = \mathbf{y}$, where the $m \times n$ matrix $\mathbf{A}$ and the $1 \times n$ vector $\mathbf{y}$ are given, and the $1 \times m$ vector $\mathbf{x}$ is to be solved. Applying Theorems 3.1 and 3.2 to the transpose of the equation, we can readily obtain the following result.

Corollary 3.2

1. Given an $m \times n$ matrix $\mathbf{A}$, a solution $\mathbf{x}$ exists in $\mathbf{xA} = \mathbf{y}$, for any $\mathbf{y}$, if and only if $\mathbf{A}$ has full column rank.

2. Given an $m \times n$ matrix $\mathbf{A}$ and an $1 \times n$ vector $\mathbf{y}$, let $\mathbf{x}_p$ be a solution of $\mathbf{xA} = \mathbf{y}$ and let $k = m - \rho(\mathbf{A})$. If $k = 0$, the solution $\mathbf{x}_p$ is unique. If $k > 0$, then for any α_i, $i = 1, 2, \ldots, k$, the vector

$$\mathbf{x} = \mathbf{x}_p + \alpha_1 \mathbf{n}_1 + \cdots + \alpha_k \mathbf{n}_k$$

is a solution of $\mathbf{xA} = \mathbf{y}$, where $\mathbf{n}_i \mathbf{A} = \mathbf{0}$ and the set $\{\mathbf{n}_1, \ldots, \mathbf{n}_k\}$ is linearly independent.

In MATLAB, the solution of $\mathbf{Ax} = \mathbf{y}$ can be obtained by typing A\y. Note the use of backslash, which denotes matrix left division. For example, for the equation in (3.16), typing

```
a=[0 1 1 2;1 2 3 4;2 0 2 0];y=[-4;-8;0];
a\y
```

yields `[0 -4 0 0]'`. The solution of $\mathbf{xA} = \mathbf{y}$ can be obtained by typing $\mathbf{y}/\mathbf{A}$. Here we use slash, which denotes matrix right division.

Determinant and inverse of square matrices The rank of a matrix is defined as the number of linearly independent columns or rows. It can also be defined using the determinant. The determinant of a 1×1 matrix is defined as itself. For $n = 2, 3, \ldots$, the determinant of $n \times n$ square matrix $\mathbf{A} = [a_{ij}]$ is defined recursively as, for any chosen j,

$$\det \mathbf{A} = \sum_{i}^{n} a_{ij} c_{ij} \tag{3.18}$$

where a_{ij} denotes the entry at the ith row and jth column of $\mathbf{A}$. Equation (3.18) is called the *Laplace expansion*. The number c_{ij} is the *cofactor* corresponding to a_{ij} and equals $(-1)^{i+j} \det M_{ij}$, where M_{ij} is the $(n-1) \times (n-1)$ submatrix of $\mathbf{A}$ by deleting its ith row and jth column. If $\mathbf{A}$ is diagonal or triangular, then $\det \mathbf{A}$ equals the product of all diagonal entries.

The determinant of any $r \times r$ submatrix of $\mathbf{A}$ is called a *minor* of order r. Then the rank can be defined as the largest order of all nonzero minors of $\mathbf{A}$. In other words, if $\mathbf{A}$ has rank r, then there is at least one nonzero minor of order r, and every minor of order larger than r is zero. A square matrix is said to be *nonsingular* if its determinant is nonzero. Thus a nonsingular square matrix has full rank and all its columns (rows) are linearly independent.

The *inverse* of a nonsingular square matrix $\mathbf{A} = [a_{ij}]$ is denoted by $\mathbf{A}^{-1}$. The inverse has the property $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ and can be computed as

$$\mathbf{A}^{-1} = \frac{\text{Adj } \mathbf{A}}{\det \mathbf{A}} = \frac{1}{\det \mathbf{A}} [c_{ij}]' \tag{3.19}$$

where c_{ij} is the cofactor. If a matrix is singular, its inverse does not exist. If $\mathbf{A}$ is 2×2, then we have

$$\mathbf{A}^{-1} := \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \tag{3.20}$$

Thus the inverse of a 2×2 matrix is very simple: interchanging diagonal entries and changing the sign of off-diagonal entries (without changing position) and dividing the resulting matrix by the determinant of $\mathbf{A}$. In general, using (3.19) to compute the inverse is complicated. If $\mathbf{A}$ is triangular, it is simpler to compute its inverse by solving $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$. Note that the inverse of a triangular matrix is again triangular. The MATLAB function `inv` computes the inverse of $\mathbf{A}$.

Theorem 3.3

Consider $\mathbf{Ax} = \mathbf{y}$ with $\mathbf{A}$ square.

1. If $\mathbf{A}$ is nonsingular, then the equation has a unique solution for every $\mathbf{y}$ and the solution equals $\mathbf{A}^{-1}\mathbf{y}$. In particular, the only solution of $\mathbf{Ax} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$.

2. The homogeneous equation $\mathbf{Ax} = \mathbf{0}$ has nonzero solutions if and only if $\mathbf{A}$ is singular. The number of linearly independent solutions equals the nullity of $\mathbf{A}$.

3.4 Similarity Transformation

Consider an $n \times n$ matrix $\mathbf{A}$. It maps $\mathcal{R}^n$ into itself. If we associate with $\mathcal{R}^n$ the orthonormal basis $\{\mathbf{i}_1, \mathbf{i}_2, \ldots, \mathbf{i}_n\}$ in (3.8), then the ith column of $\mathbf{A}$ is the representation of $\mathbf{Ai}_i$ with respect to the orthonormal basis. Now if we select a different set of basis $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$, then the matrix $\mathbf{A}$ has a different representation $\bar{\mathbf{A}}$. It turns out that *the ith column of $\bar{\mathbf{A}}$ is the representation of $\mathbf{Aq}_i$ with respect to the basis $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$*. This is illustrated by the example that follows.

Example 3.4 Consider the matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & -1 \\ -2 & 1 & 0 \\ 4 & 3 & 1 \end{bmatrix} \tag{3.21}$$

Let $\mathbf{b} = [0\ 0\ 1]'$. Then we have

$$\mathbf{Ab} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{A}^2\mathbf{b} = \mathbf{A}(\mathbf{Ab}) = \begin{bmatrix} -4 \\ 2 \\ -3 \end{bmatrix}, \quad \mathbf{A}^3\mathbf{b} = \mathbf{A}(\mathbf{A}^2\mathbf{b}) = \begin{bmatrix} -5 \\ 10 \\ -13 \end{bmatrix}$$

It can be verified that the following relation holds:

$$\mathbf{A}^3\mathbf{b} = 17\mathbf{b} - 15\mathbf{Ab} + 5\mathbf{A}^2\mathbf{b} \tag{3.22}$$

Because the three vectors $\mathbf{b}$, $\mathbf{Ab}$, and $\mathbf{A}^2\mathbf{b}$ are linearly independent, they can be used as a basis. We now compute the representation of $\mathbf{A}$ with respect to the basis. It is clear that

$$\mathbf{A}(\mathbf{b}) = [\mathbf{b}\ \ \mathbf{Ab}\ \ \mathbf{A}^2\mathbf{b}] \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{A}(\mathbf{Ab}) = [\mathbf{b}\ \ \mathbf{Ab}\ \ \mathbf{A}^2\mathbf{b}] \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\mathbf{A}(\mathbf{A}^2\mathbf{b}) = [\mathbf{b}\ \ \mathbf{Ab}\ \ \mathbf{A}^2\mathbf{b}] \begin{bmatrix} 17 \\ -15 \\ 5 \end{bmatrix}$$

where the last equation is obtained from (3.22). Thus the representation of $\mathbf{A}$ with respect to the basis $\{\mathbf{b}, \mathbf{Ab}, \mathbf{A}^2\mathbf{b}\}$ is

$$\bar{\mathbf{A}} = \begin{bmatrix} 0 & 0 & 17 \\ 1 & 0 & -15 \\ 0 & 1 & 5 \end{bmatrix} \tag{3.23}$$

The preceding discussion can be extended to the general case. Let $\mathbf{A}$ be an $n \times n$ matrix. If there exists an $n \times 1$ vector $\mathbf{b}$ such that the n vectors $\mathbf{b}, \mathbf{Ab}, \ldots, \mathbf{A}^{n-1}\mathbf{b}$ are linearly independent and if

$$\mathbf{A}^n = \beta_1 \mathbf{b} + \beta_2 \mathbf{Ab} + \cdots + \beta_n \mathbf{A}^{n-1}\mathbf{b}$$

then the representation of $\mathbf{A}$ with respect to the basis $\{\mathbf{b},\ \mathbf{Ab},\ \ldots,\ \mathbf{A}^{n-1}\mathbf{b}\}$ is

$$\bar{\mathbf{A}} = \begin{bmatrix} 0 & 0 & \cdots & 0 & \beta_1 \\ 1 & 0 & \cdots & 0 & \beta_2 \\ 0 & 1 & \cdots & 0 & \beta_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \beta_{n-1} \\ 0 & 0 & \cdots & 1 & \beta_n \end{bmatrix} \tag{3.24}$$

This matrix is said to be in a *companion* form.

Consider the equation

$$\mathbf{Ax} = \mathbf{y} \tag{3.25}$$

The square matrix $\mathbf{A}$ maps $\mathbf{x}$ in $\mathcal{R}^n$ into $\mathbf{y}$ in $\mathcal{R}^n$. With respect to the basis $\{\mathbf{q}_1,\ \mathbf{q}_2,\ \ldots,\ \mathbf{q}_n\}$, the equation becomes

$$\bar{\mathbf{A}}\bar{\mathbf{x}} = \bar{\mathbf{y}} \tag{3.26}$$

where $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ are the representations of $\mathbf{x}$ and $\mathbf{y}$ with respect to the basis $\{\mathbf{q}_1,\ \mathbf{q}_2,\ \ldots,\ \mathbf{q}_n\}$. As discussed in (3.7), they are related by

$$\mathbf{x} = \mathbf{Q}\bar{\mathbf{x}} \qquad \mathbf{y} = \mathbf{Q}\bar{\mathbf{y}}$$

with

$$\mathbf{Q} = [\mathbf{q}_1\ \ \mathbf{q}_2\ \ \cdots\ \ \mathbf{q}_n] \tag{3.27}$$

an $n \times n$ nonsingular matrix. Substituting these into (3.25) yields

$$\mathbf{AQ}\bar{\mathbf{x}} = \mathbf{Q}\bar{\mathbf{y}} \quad \text{or} \quad \mathbf{Q}^{-1}\mathbf{AQ}\bar{\mathbf{x}} = \bar{\mathbf{y}} \tag{3.28}$$

Comparing this with (3.26) yields

$$\bar{\mathbf{A}} = \mathbf{Q}^{-1}\mathbf{AQ} \quad \text{or} \quad \mathbf{A} = \mathbf{Q}\bar{\mathbf{A}}\mathbf{Q}^{-1} \tag{3.29}$$

This is called the *similarity transformation* and $\mathbf{A}$ and $\bar{\mathbf{A}}$ are said to be *similar*. We write (3.29) as

$$\mathbf{AQ} = \mathbf{Q}\bar{\mathbf{A}}$$

or

$$\mathbf{A}[\mathbf{q}_1\ \ \mathbf{q}_2\ \ \cdots\ \ \mathbf{q}_n] = [\mathbf{Aq}_1\ \ \mathbf{Aq}_2\ \ \cdots\ \ \mathbf{Aq}_n] = [\mathbf{q}_1\ \ \mathbf{q}_2\ \ \cdots\ \ \mathbf{q}_n]\bar{\mathbf{A}}$$

This shows that the ith column of $\bar{\mathbf{A}}$ is indeed the representation of $\mathbf{Aq}_i$ with respect to the basis $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$.

3.5 Diagonal Form and Jordan Form

A square matrix $\mathbf{A}$ has different representations with respect to different sets of basis. In this section, we introduce a set of basis so that the representation will be diagonal or block diagonal.

A real or complex number λ is called an *eigenvalue* of the $n \times n$ real matrix $\mathbf{A}$ if there exists a nonzero vector $\mathbf{x}$ such that $\mathbf{Ax} = \lambda\mathbf{x}$. Any nonzero vector $\mathbf{x}$ satisfying $\mathbf{Ax} = \lambda\mathbf{x}$ is called a (right) *eigenvector* of $\mathbf{A}$ associated with eigenvalue λ. In order to find the eigenvalue of $\mathbf{A}$, we write $\mathbf{Ax} = \lambda\mathbf{x} = \lambda\mathbf{Ix}$ as

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \tag{3.30}$$

where $\mathbf{I}$ is the unit matrix of order n. This is a homogeneous equation. If the matrix $(\mathbf{A} - \lambda\mathbf{I})$ is nonsingular, then the only solution of (3.30) is $\mathbf{x} = \mathbf{0}$ (Theorem 3.3). Thus in order for (3.30) to have a nonzero solution $\mathbf{x}$, the matrix $(\mathbf{A} - \lambda\mathbf{I})$ must be singular or have determinant 0. We define

$$\Delta(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A})$$

It is a monic polynomial of degree n with real coefficients and is called the *characteristic polynomial* of $\mathbf{A}$. A polynomial is called monic if its leading coefficient is 1. If λ is a root of the characteristic polynomial, then the determinant of $(\mathbf{A} - \lambda\mathbf{I})$ is 0 and (3.30) has at least one nonzero solution. Thus every root of $\Delta(\lambda)$ is an eigenvalue of $\mathbf{A}$. Because $\Delta(\lambda)$ has degree n, the $n \times n$ matrix $\mathbf{A}$ has n eigenvalues (not necessarily all distinct).

We mention that the matrices

$$\begin{bmatrix} 0 & 0 & 0 & -\alpha_4 \\ 1 & 0 & 0 & -\alpha_3 \\ 0 & 1 & 0 & -\alpha_2 \\ 0 & 0 & 1 & -\alpha_1 \end{bmatrix} \qquad \begin{bmatrix} -\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

and their transposes

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\alpha_4 & -\alpha_3 & -\alpha_2 & -\alpha_1 \end{bmatrix} \qquad \begin{bmatrix} -\alpha_1 & 1 & 0 & 0 \\ -\alpha_2 & 0 & 1 & 0 \\ -\alpha_3 & 0 & 0 & 1 \\ -\alpha_4 & 0 & 0 & 0 \end{bmatrix}$$

all have the following characteristic polynomial:

$$\Delta(\lambda) = \lambda^4 + \alpha_1\lambda^3 + \alpha_2\lambda^2 + \alpha_3\lambda + \alpha_4$$

These matrices can easily be formed from the coefficients of $\Delta(\lambda)$ and are called *companion-form* matrices. The companion-form matrices will arise repeatedly later. The matrix in (3.24) is in such a form.

Eigenvalues of A are all distinct Let λ_i, $i = 1, 2, \ldots, n$, be the eigenvalues of $\mathbf{A}$ and be all distinct. Let $\mathbf{q}_i$ be an eigenvector of $\mathbf{A}$ associated with λ_i; that is, $\mathbf{A}\mathbf{q}_i = \lambda_i \mathbf{q}_i$. Then the set of eigenvectors $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ is linearly independent and can be used as a basis. Let $\hat{\mathbf{A}}$ be the representation of $\mathbf{A}$ with respect to this basis. Then the first column of $\hat{\mathbf{A}}$ is the representation of $\mathbf{A}\mathbf{q}_1 = \lambda_1 \mathbf{q}_1$ with respect to $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$. From

$$\mathbf{A}\mathbf{q}_1 = \lambda_1 \mathbf{q}_1 = [\mathbf{q}_1 \ \ \mathbf{q}_2 \ \ \cdots \ \ \mathbf{q}_n] \begin{bmatrix} \lambda_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

we conclude that the first column of $\hat{\mathbf{A}}$ is $[\lambda_1 \ 0 \ \cdots \ 0]'$. The second column of $\hat{\mathbf{A}}$ is the representation of $\mathbf{A}\mathbf{q}_2 = \lambda_2 \mathbf{q}_2$ with respect to $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$, that is, $[0 \ \lambda_1 \ 0 \ \cdots \ 0]'$. Proceeding forward, we can establish

$$\hat{\mathbf{A}} = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{3.31}$$

This is a diagonal matrix. Thus we conclude that every matrix with distinct eigenvalues has a diagonal matrix representation by using its eigenvectors as a basis. Different orderings of eigenvectors will yield different diagonal matrices for the same $\mathbf{A}$.

If we define

$$\mathbf{Q} = [\mathbf{q}_1 \ \ \mathbf{q}_2 \ \ \cdots \ \ \mathbf{q}_n] \tag{3.32}$$

then the matrix $\hat{\mathbf{A}}$ equals

$$\hat{\mathbf{A}} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} \tag{3.33}$$

as derived in (3.29). Computing (3.33) by hand is not simple because of the need to compute the inverse of $\mathbf{Q}$. However, if we know $\hat{\mathbf{A}}$, then we can verify (3.33) by checking $\mathbf{Q}\hat{\mathbf{A}} = \mathbf{A}\mathbf{Q}$.

Example 3.5 Consider the matrix

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

Its characteristic polynomial is

$$\begin{aligned} \Delta(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A}) &= \det \begin{bmatrix} \lambda & 0 & 0 \\ -1 & \lambda & -2 \\ 0 & -1 & \lambda - 1 \end{bmatrix} \\ &= \lambda[\lambda(\lambda - 1) - 2] = (\lambda - 2)(\lambda + 1)\lambda \end{aligned}$$

Thus $\mathbf{A}$ has eigenvalues 2, -1, and 0. The eignevector associated with $\lambda = 2$ is any nonzero solution of

$$(\mathbf{A} - 2\mathbf{I})\mathbf{q}_1 = \begin{bmatrix} -2 & 0 & 0 \\ 1 & -2 & 2 \\ 0 & 1 & -1 \end{bmatrix} \mathbf{q}_1 = \mathbf{0}$$

Thus $\mathbf{q}_1 = [0 \;\; 1 \;\; 1]'$ is an eigenvector associated with $\lambda = 2$. Note that the eigenvector is not unique, $[0 \;\; \alpha \;\; \alpha]'$ for any nonzero real α can also be chosen as an eigenvector. The eigenvector associated with $\lambda = -1$ is any nonzero solution of

$$(\mathbf{A} - (-1)\mathbf{I})\mathbf{q}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \mathbf{q}_2 = \mathbf{0}$$

which yields $\mathbf{q}_2 = [0 \;\; -2 \;\; 1]'$. Similarly, the eigenvector associated with $\lambda = 0$ can be computed as $\mathbf{q}_3 = [2 \;\; 1 \;\; -1]'$. Thus the representation of $\mathbf{A}$ with respect to $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ is

$$\hat{\mathbf{A}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.34}$$

It is a diagonal matrix with eigenvalues on the diagonal. This matrix can also be obtained by computing

$$\hat{\mathbf{A}} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$$

with

$$\mathbf{Q} = [\mathbf{q}_1 \; \mathbf{q}_2 \; \mathbf{q}_3] = \begin{bmatrix} 0 & 0 & 2 \\ 1 & -2 & 1 \\ 1 & 1 & -1 \end{bmatrix} \tag{3.35}$$

However, it is simpler to verify $\mathbf{Q}\hat{\mathbf{A}} = \mathbf{A}\mathbf{Q}$ or

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

The result in this example can easily be obtained using MATLAB. Typing

```
a=[0 0 0;1 0 2;0 1 1];  [q,d]=eig(a)
```

yields

$$q = \begin{bmatrix} 0 & 0 & 0.8186 \\ 0.7071 & 0.8944 & 0.4082 \\ 0.7071 & -0.4472 & -0.4082 \end{bmatrix} \qquad d = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

where `d` is the diagonal matrix in (3.34). The matrix `q` is different from the $\mathbf{Q}$ in (3.35); but their corresponding columns differ only by a constant. This is due to nonuniqueness of eigenvectors and every column of `q` is normalized to have norm 1 in MATLAB. If we type `eig(a)` without the left-hand-side argument, then MATLAB generates only the three eigenvalues `2, -1, 0`.

We mention that eigenvalues in MATLAB are *not* computed from the characteristic polynomial. Computing the characteristic polynomial using the Laplace expansion and then computing its roots are not numerically reliable, especially when there are repeated roots. Eigenvalues are computed in MATLAB directly from the matrix by using similarity transformations. Once all eigenvalues are computed, the characteristic polynomial equals $\prod(\lambda - \lambda_i)$. In MATLAB, typing `r=eig(a);poly(r)` yields the characteristic polynomial.

Example 3.6 Consider the matrix

$$\mathbf{A} = \begin{bmatrix} -1 & 1 & 1 \\ 0 & 4 & -13 \\ 0 & 1 & 0 \end{bmatrix}$$

Its characteristic polynomial is $(\lambda+1)(\lambda^2-4\lambda+13)$. Thus $\mathbf{A}$ has eigenvalues $-1,\ 2\pm 3j$. Note that complex conjugate eigenvalues must appear in pairs because $\mathbf{A}$ has only real coefficients. The eigenvectors associated with -1 and $2+3j$ are, respectively, $[1\ \ 0\ \ 0]'$ and $[j\ \ -3+2j\ \ j]'$. The eigenvector associated with $\lambda = 2-3j$ is $[-j\ \ -3-2j\ \ -j]'$, the complex conjugate of the eigenvector associated with $\lambda = 2+3j$. Thus we have

$$\mathbf{Q} = \begin{bmatrix} 1 & j & -j \\ 0 & -3+2j & -3-2j \\ 0 & j & j \end{bmatrix} \quad \text{and} \quad \hat{\mathbf{A}} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2+3j & 0 \\ 0 & 0 & 2-3j \end{bmatrix} \tag{3.36}$$

The MATLAB function `[q,d]=eig(a)` yields

$$q = \begin{bmatrix} 1 & 0.2582j & -0.2582j \\ 0 & -0.7746+0.5164j & -0.7746-0.5164j \\ 0 & 0.2582j & -0.2582j \end{bmatrix}$$

and

$$d = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2+3j & 0 \\ 0 & 0 & 2-3j \end{bmatrix}$$

All discussion in the preceding example applies here.

Complex eigenvalues Even though the data we encounter in practice are all real numbers, complex numbers may arise when we compute eigenvalues and eigenvectors. To deal with this problem, we must extend real linear spaces into complex linear spaces and permit all scalars such as α_i in (3.4) to assume complex numbers. To see the reason, we consider

$$\mathbf{Av} = \begin{bmatrix} 1 & 1+j \\ 1-j & 2 \end{bmatrix} \mathbf{v} = \mathbf{0} \tag{3.37}$$

If we restrict $\mathbf{v}$ to real vectors, then (3.37) has no nonzero solution and the two columns of $\mathbf{A}$ are linearly independent. However, if $\mathbf{v}$ is permitted to assume complex numbers, then $\mathbf{v} = [-2\ \ 1-j]'$ is a nonzero solution of (3.37). Thus the two columns of $\mathbf{A}$ are linearly dependent and $\mathbf{A}$ has rank 1. This is the rank obtained in MATLAB. Therefore, whenever complex eigenvalues arise, we consider complex linear spaces and complex scalars and transpose is replaced by complex-conjugate transpose. By so doing, all concepts and results developed for real vectors and matrices can be applied to complex vectors and matrices. Incidentally, the diagonal matrix with complex eigenvalues in (3.36) can be transformed into a very useful real matrix as we will discuss in Section 4.3.1.

Eigenvalues of A are not all distinct An eigenvalue with multiplicity 2 or higher is called a *repeated* eigenvalue. In contrast, an eigenvalue with multiplicity 1 is called a *simple* eigenvalue. If **A** has only simple eigenvalues, it always has a diagonal-form representation. If **A** has repeated eigenvalues, then it may not have a diagonal form representation. However, it has a block-diagonal and triangular-form representation as we will discuss next.

Consider an $n \times n$ matrix **A** with eigenvalue λ and multiplicity n. In other words, **A** has only one distinct eigenvalue. To simplify the discussion, we assume $n = 4$. Suppose the matrix $(\mathbf{A} - \lambda\mathbf{I})$ has rank $n - 1 = 3$ or, equivalently, nullity 1; then the equation

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{q} = \mathbf{0}$$

has only one independent solution. Thus **A** has only one eigenvector associated with λ. We need $n - 1 = 3$ more linearly independent vectors to form a basis for $\mathcal{R}^4$. The three vectors $\mathbf{q}_2$, $\mathbf{q}_3$, $\mathbf{q}_4$ will be chosen to have the properties $(\mathbf{A} - \lambda\mathbf{I})^2\mathbf{q}_2 = \mathbf{0}$, $(\mathbf{A} - \lambda\mathbf{I})^3\mathbf{q}_3 = \mathbf{0}$, and $(\mathbf{A} - \lambda\mathbf{I})^4\mathbf{q}_4 = \mathbf{0}$.

A vector **v** is called a *generalized eigenvector* of grade n if

$$(\mathbf{A} - \lambda\mathbf{I})^n\mathbf{v} = \mathbf{0}$$

and

$$(\mathbf{A} - \lambda\mathbf{I})^{n-1}\mathbf{v} \neq \mathbf{0}$$

If $n = 1$, they reduce to $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$ and $\mathbf{v} \neq \mathbf{0}$ and **v** is an ordinary eigenvector. For $n = 4$, we define

$$\begin{aligned}
\mathbf{v}_4 &:= \mathbf{v} \\
\mathbf{v}_3 &:= (\mathbf{A} - \lambda\mathbf{I})\mathbf{v}_4 = (\mathbf{A} - \lambda\mathbf{I})\mathbf{v} \\
\mathbf{v}_2 &:= (\mathbf{A} - \lambda\mathbf{I})\mathbf{v}_3 = (\mathbf{A} - \lambda\mathbf{I})^2\mathbf{v} \\
\mathbf{v}_1 &:= (\mathbf{A} - \lambda\mathbf{I})\mathbf{v}_2 = (\mathbf{A} - \lambda\mathbf{I})^3\mathbf{v}
\end{aligned}$$

They are called a chain of generalized eigenvectors of length $n = 4$ and have the properties $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v}_1 = \mathbf{0}$, $(\mathbf{A} - \lambda\mathbf{I})^2\mathbf{v}_2 = \mathbf{0}$, $(\mathbf{A} - \lambda\mathbf{I})^3\mathbf{v}_3 = \mathbf{0}$, and $(\mathbf{A} - \lambda\mathbf{I})^4\mathbf{v}_4 = \mathbf{0}$. These vectors, as generated, are automatically linearly independent and can be used as a basis. From these equations, we can readily obtain

$$\begin{aligned}
\mathbf{A}\mathbf{v}_1 &= \lambda\mathbf{v}_1 \\
\mathbf{A}\mathbf{v}_2 &= \mathbf{v}_1 + \lambda\mathbf{v}_2 \\
\mathbf{A}\mathbf{v}_3 &= \mathbf{v}_2 + \lambda\mathbf{v}_3 \\
\mathbf{A}\mathbf{v}_4 &= \mathbf{v}_3 + \lambda\mathbf{v}_4
\end{aligned}$$

Then the representation of $\mathbf{A}$ with respect to the basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$ is

$$\mathbf{J} := \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{bmatrix} \tag{3.38}$$

We verify this for the first and last columns. The first column of $\mathbf{J}$ is the representation of $\mathbf{A}\mathbf{v}_1 = \lambda\mathbf{v}_1$ with respect to $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$, which is $[\lambda\ \ 0\ \ 0\ \ 0]'$. The last column of $\mathbf{J}$ is the representation of $\mathbf{A}\mathbf{v}_4 = \mathbf{v}_3 + \lambda\mathbf{v}_4$ with respect to $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \mathbf{v}_4\}$, which is $[0\ \ 0\ \ 1\ \ \lambda]'$. This verifies the representation in (3.38). The matrix $\mathbf{J}$ has eigenvalues on the diagonal and 1 on the superdiagonal. If we reverse the order of the basis, then the 1's will appear on the subdiagonal. The matrix is called a *Jordan* block of order $n = 4$.

If $(\mathbf{A} - \lambda\mathbf{I})$ has rank $n - 2$ or, equivalently, nullity 2, then the equation

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{q} = \mathbf{0}$$

has two linearly independent solutions. Thus $\mathbf{A}$ has two linearly independent eigenvectors and we need $(n - 2)$ generalized eigenvectors. In this case, there exist two chains of generalized eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ and $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_l\}$ with $k + l = n$. If $\mathbf{v}_1$ and $\mathbf{u}_1$ are linearly independent, then the set of n vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{u}_1, \ldots, \mathbf{u}_l\}$ is linearly independent and can be used as a basis. With respect to this basis, the representation of $\mathbf{A}$ is a block diagonal matrix of form

$$\hat{\mathbf{A}} = \text{diag}\{\mathbf{J}_1, \mathbf{J}_2\}$$

where $\mathbf{J}_1$ and $\mathbf{J}_2$ are, respectively, Jordan blocks of order k and l.

Now we discuss a specific example. Consider a 5×5 matrix $\mathbf{A}$ with repeated eigenvalue λ_1 with multiplicity 4 and simple eigenvalue λ_2. Then there exists a nonsingular matrix $\mathbf{Q}$ such that

$$\hat{\mathbf{A}} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$$

assumes one of the following forms

$$\hat{\mathbf{A}}_1 = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix} \qquad \hat{\mathbf{A}}_2 = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix}$$

$$\hat{\mathbf{A}}_3 = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix} \qquad \hat{\mathbf{A}}_4 = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix}$$

$$\hat{\mathbf{A}}_5 = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix} \tag{3.39}$$

The first matrix occurs when the nullity of $(\mathbf{A} - \lambda_1 \mathbf{I})$ is 1. If the nullity is 2, then $\hat{\mathbf{A}}$ has two Jordan blocks associated with λ_1; it may assume the form in $\hat{\mathbf{A}}_2$ or in $\hat{\mathbf{A}}_3$. If $(\mathbf{A} - \lambda_1 \mathbf{I})$ has nullity 3, then $\hat{\mathbf{A}}$ has three Jordan blocks associated with λ_1 as shown in $\hat{\mathbf{A}}_4$. Certainly, the positions of the Jordan blocks can be changed by changing the order of the basis. If the nullity is 4, then $\hat{\mathbf{A}}$ is a diagonal matrix as shown in $\hat{\mathbf{A}}_5$. All these matrices are triangular and block diagonal with Jordan blocks on the diagonal; they are said to be in Jordan form. A diagonal matrix is a degenerated Jordan form; its Jordan blocks all have order 1. If $\mathbf{A}$ can be diagonalized, we can use `[q,d]=eig(a)` to generate $\mathbf{Q}$ and $\hat{\mathbf{A}}$ as shown in Examples 3.5 and 3.6. If $\mathbf{A}$ cannot be diagonized, $\mathbf{A}$ is said to be *defective* and `[q,d]=eig(a)` will yield an incorrect solution. In this case, we may use the MATLAB function `[q,d]=jordan(a)`. However, `jordan` will yield a correct result only if $\mathbf{A}$ has integers or ratios of small integers as its entries.

Jordan-form matrices are triangular and block diagonal and can be used to establish many general properties of matrices. For example, because $\det(\mathbf{CD}) = \det \mathbf{C} \det \mathbf{D}$ and $\det \mathbf{Q} \det \mathbf{Q}^{-1} = \det \mathbf{I} = 1$, from $\mathbf{A} = \mathbf{Q}\hat{\mathbf{A}}\mathbf{Q}^{-1}$, we have

$$\det \mathbf{A} = \det \mathbf{Q} \det \hat{\mathbf{A}} \det \mathbf{Q}^{-1} = \det \hat{\mathbf{A}}$$

The determinant of $\hat{\mathbf{A}}$ is the product of all diagonal entries or, equivalently, all eigenvalues of $\mathbf{A}$. Thus we have

$$\det \mathbf{A} = \text{product of all eigenvalues of } \mathbf{A}$$

which implies that $\mathbf{A}$ *is nonsingular if and only if it has no zero eigenvalue.*

We discuss a useful property of Jordan blocks to conclude this section. Consider the Jordan block in (3.38) with order 4. Then we have

$$(\mathbf{J} - \lambda \mathbf{I}) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (\mathbf{J} - \lambda \mathbf{I})^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$(\mathbf{J} - \lambda \mathbf{I})^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.40}$$

and $(\mathbf{J} - \lambda \mathbf{I})^k = \mathbf{0}$ for $k \geq 4$. This is called *nilpotent*.

3.6 Functions of a Square Matrix

This section studies functions of a square matrix. We use Jordan form extensively because many properties of functions can almost be visualized in terms of Jordan form. We study first polynomials and then general functions of a square matrix.

Polynomials of a square matrix Let $\mathbf{A}$ be a square matrix. If k is a positive integer, we define

$$\mathbf{A}^k := \mathbf{A}\mathbf{A} \cdots \mathbf{A} \quad (k \text{ terms})$$

and $\mathbf{A}^0 = \mathbf{I}$. Let $f(\lambda)$ be a polynomial such as $f(\lambda) = \lambda^3 + 2\lambda^2 - 6$ or $(\lambda + 2)(4\lambda - 3)$. Then $f(\mathbf{A})$ is defined as

$$f(\mathbf{A}) = \mathbf{A}^3 + 2\mathbf{A}^2 - 6\mathbf{I} \quad \text{or} \quad f(\mathbf{A}) = (\mathbf{A} + 2\mathbf{I})(4\mathbf{A} - 3\mathbf{I})$$

If $\mathbf{A}$ is block diagonal, such as

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix}$$

where $\mathbf{A}_1$ and $\mathbf{A}_2$ are square matrices of any order, then it is straightforward to verify

$$\mathbf{A}^k = \begin{bmatrix} \mathbf{A}_1^k & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2^k \end{bmatrix} \quad \text{and} \quad f(\mathbf{A}) = \begin{bmatrix} f(\mathbf{A}_1) & \mathbf{0} \\ \mathbf{0} & f(\mathbf{A}_2) \end{bmatrix} \tag{3.41}$$

Consider the similarity transformation $\hat{\mathbf{A}} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$ or $\mathbf{A} = \mathbf{Q}\hat{\mathbf{A}}\mathbf{Q}^{-1}$. Because

$$\mathbf{A}^k = (\mathbf{Q}\hat{\mathbf{A}}\mathbf{Q}^{-1})(\mathbf{Q}\hat{\mathbf{A}}\mathbf{Q}^{-1})\cdots(\mathbf{Q}\hat{\mathbf{A}}\mathbf{Q}^{-1}) = \mathbf{Q}\hat{\mathbf{A}}^k\mathbf{Q}^{-1}$$

we have

$$f(\mathbf{A}) = \mathbf{Q}f(\hat{\mathbf{A}})\mathbf{Q}^{-1} \quad \text{or} \quad f(\hat{\mathbf{A}}) = \mathbf{Q}^{-1}f(\mathbf{A})\mathbf{Q} \tag{3.42}$$

A *monic* polynomial is a polynomial with 1 as its leading coefficient. The *minimal polynomial* of $\mathbf{A}$ is defined as the monic polynomial $\psi(\lambda)$ of least degree such that $\psi(\mathbf{A}) = \mathbf{0}$. Note that the $\mathbf{0}$ is a zero matrix of the same order as $\mathbf{A}$. A direct consequence of (3.42) is that $f(\mathbf{A}) = \mathbf{0}$ if and only if $f(\hat{\mathbf{A}}) = \mathbf{0}$. Thus $\mathbf{A}$ and $\hat{\mathbf{A}}$ or, more general, all similar matrices have the same minimal polynomial. Computing the minimal polynomial directly from $\mathbf{A}$ is not simple (see Problem 3.25); however, if the Jordan-form representation of $\mathbf{A}$ is available, the minimal polynomial can be read out by inspection.

Let λ_i be an eigenvalue of $\mathbf{A}$ with multiplicity n_i. That is, the characteristic polynomial of $\mathbf{A}$ is

$$\Delta(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A}) = \prod_i (\lambda - \lambda_i)^{n_i}$$

Suppose the Jordan form of $\mathbf{A}$ is known. Associated with each eigenvalue, there may be one or more Jordan blocks. The *index* of λ_i, denoted by $\bar{n}_i$, is defined as the largest order of all Jordan blocks associated with λ_i. Clearly we have $\bar{n}_i \leq n_i$. For example, the multiplicities of λ_1 in all five matrices in (3.39) are 4; their indices are, respectively, 4, 3, 2, 2, and 1. The multiplicities and indices of λ_2 in all five matrices in (3.39) are all 1. Using the indices of all eigenvalues, the minimal polynomial can be expressed as

$$\psi(\lambda) = \prod_i (\lambda - \lambda_i)^{\bar{n}_i}$$

with degree $\bar{n} = \sum \bar{n}_i \leq \sum n_i = n$ = dimension of $\mathbf{A}$. For example, the minimal polynomials of the five matrices in (3.39) are

$$\psi_1 = (\lambda - \lambda_1)^4(\lambda - \lambda_2) \qquad \psi_2 = (\lambda - \lambda_1)^3(\lambda - \lambda_2)$$
$$\psi_3 = (\lambda - \lambda_1)^2(\lambda - \lambda_2) \qquad \psi_4 = (\lambda - \lambda_1)^2(\lambda - \lambda_2)$$
$$\psi_5 = (\lambda - \lambda_1)(\lambda - \lambda_2)$$

Their characteristic polynomials, however, all equal

$$\Delta(\lambda) = (\lambda - \lambda_1)^4(\lambda - \lambda_2)$$

We see that the minimal polynomial is a factor of the characteristic polynomial and has a degree less than or equal to the degree of the characteristic polynomial. Clearly, if all eigenvalues of $\mathbf{A}$ are distinct, then the minimal polynomial equals the characteristic polynomial.

Using the nilpotent property in (3.40), we can show that

$$\psi(\mathbf{A}) = \mathbf{0}$$

and that no polynomial of lesser degree meets the condition. Thus $\psi(\lambda)$ as defined is the minimal polynomial.

Theorem 3.4 (Cayley-Hamilton theorem)

Let

$$\Delta(\lambda) = \det(\lambda\mathbf{I} - \mathbf{A}) = \lambda^n + \alpha_1\lambda^{n-1} + \cdots + \alpha_{n-1}\lambda + \alpha_n$$

be the characteristic polynomial of $\mathbf{A}$. Then

$$\Delta(\mathbf{A}) = \mathbf{A}^n + \alpha_1\mathbf{A}^{n-1} + \cdots + \alpha_{n-1}\mathbf{A} + \alpha_n\mathbf{I} = \mathbf{0} \tag{3.43}$$

In words, a matrix satisfies its own characteristic polynomial. Because $n_i \geq \bar{n}_i$, the characteristic polynomial contains the minimal polynomial as a factor or $\Delta(\lambda) = \psi(\lambda)h(\lambda)$ for some polynomial $h(\lambda)$. Because $\psi(\mathbf{A}) = \mathbf{0}$, we have $\Delta(\mathbf{A}) = \psi(\mathbf{A})h(\mathbf{A}) = \mathbf{0} \cdot h(\mathbf{A}) = \mathbf{0}$. This establishes the theorem. The Cayley–Hamilton theorem implies that $\mathbf{A}^n$ can be written as a linear combination of $\{\mathbf{I}, \mathbf{A}, \ldots, \mathbf{A}^{n-1}\}$. Multiplying (3.43) by $\mathbf{A}$ yields

$$\mathbf{A}^{n+1} + \alpha_1\mathbf{A}^n + \cdots + \alpha_{n-1}\mathbf{A}2 + \alpha_n\mathbf{A} = \mathbf{0} \cdot \mathbf{A} = \mathbf{0}$$

which implies that $\mathbf{A}^{n+1}$ can be written as a linear combination of $\{\mathbf{A}, \mathbf{A}^2, \ldots, \mathbf{A}^n\}$, which, in turn, can be written as a linear combination of $\{\mathbf{I}, \mathbf{A}, \ldots, \mathbf{A}^{n-1}\}$. Proceeding forward, we conclude that, for any polynomial $f(\lambda)$, no matter how large its degree is, $f(\mathbf{A})$ can always be expressed as

$$f(\mathbf{A}) = \beta_0\mathbf{I} + \beta_1\mathbf{A} + \cdots + \beta_{n-1}\mathbf{A}^{n-1} \tag{3.44}$$

for some β_i. In other words, every polynomial of an $n \times n$ matrix $\mathbf{A}$ can be expressed as a linear combination of $\{\mathbf{I}, \mathbf{A}, \ldots, \mathbf{A}^{n-1}\}$. If the minimal polynomial of $\mathbf{A}$ with degree $\bar{n}$ is available, then every polynomial of $\mathbf{A}$ can be expressed as a linear combination of $\{\mathbf{I}, \mathbf{A}, \ldots, \mathbf{A}^{\bar{n}-1}\}$. This is a better result. However, because $\bar{n}$ may not be available, we discuss in the following only (3.44) with the understanding that all discussion applies to $\bar{n}$.

One way to compute (3.44) is to use long division to express $f(\lambda)$ as

$$f(\lambda) = q(\lambda)\Delta(\lambda) + h(\lambda) \tag{3.45}$$

where $q(\lambda)$ is the quotient and $h(\lambda)$ is the remainder with degree less than n. Then we have

$$f(\mathbf{A}) = q(\mathbf{A})\Delta(\mathbf{A}) + h(\mathbf{A}) = q(\mathbf{A})\mathbf{0} + h(\mathbf{A}) = h(\mathbf{A})$$

Long division is not convenient to carry out if the degree of $f(\lambda)$ is much larger than the degree of $\Delta(\lambda)$. In this case, we may solve $h(\lambda)$ directly from (3.45). Let

$$h(\lambda) := \beta_0 + \beta_1\lambda + \cdots + \beta_{n-1}\lambda^{n-1}$$

where the n unknowns β_i are to be solved. If all n eigenvalues of $\mathbf{A}$ are distinct, these β_i can be solved from the n equations

$$f(\lambda_i) = q(\lambda_i)\Delta(\lambda_i) + h(\lambda_i) = h(\lambda_i)$$

for $i = 1, 2, \ldots, n$. If $\mathbf{A}$ has repeated eigenvalues, then (3.45) must be differentiated to yield additional equations. This is stated as a theorem.

Theorem 3.5

We are given $f(\lambda)$ and an $n \times n$ matrix $\mathbf{A}$ with characteristic polynomial

$$\Delta(\lambda) = \prod_{i=1}^{m}(\lambda - \lambda_i)^{n_i}$$

where $n = \sum_{i=1}^{m} n_i$. Define

$$h(\lambda) := \beta_0 + \beta_1\lambda + \cdots + \beta_{n-1}\lambda^{n-1}$$

It is a polynomial of degree $n-1$ with n unknown coefficients. These n unknowns are to be solved from the following set of n equations:

$$f^{(l)}(\lambda_i) = h^{(l)}(\lambda_i) \quad \text{for } l = 0, 1, \ldots, n_i - 1 \quad \text{and} \quad i = 1, 2, \ldots, m$$

where

$$f^{(l)}(\lambda_i) := \left.\frac{d^l f(\lambda)}{d\lambda^l}\right|_{\lambda=\lambda_i}$$

and $h^{(l)}(\lambda_i)$ is similarly defined. Then we have

$$f(\mathbf{A}) = h(\mathbf{A})$$

and $h(\lambda)$ is said to equal $f(\lambda)$ on the spectrum of $\mathbf{A}$.

Example 3.7 Compute $\mathbf{A}^{100}$ with

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

In other words, given $f(\lambda) = \lambda^{100}$, compute $f(\mathbf{A})$. The characteristic polynomial of $\mathbf{A}$ is $\Delta(\lambda) = \lambda^2 + 2\lambda + 1 = (\lambda + 1)^2$. Let $h(\lambda) = \beta_0 + \beta_1\lambda$. On the spectrum of $\mathbf{A}$, we have

$$f(-1) = h(-1): \qquad (-1)^{100} = \beta_0 - \beta_1$$
$$f'(-1) = h'(-1): \qquad 100 \cdot (-1)^{99} = \beta_1$$

Thus we have $\beta_1 = -100$, $\beta_0 = 1 + \beta_1 = -99$, $h(\lambda) = -99 - 100\lambda$, and

$$\mathbf{A}^{100} = \beta_0\mathbf{I} + \beta_1\mathbf{A} = -99\mathbf{I} - 100\mathbf{A}$$

$$= -99\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - 100\begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} -199 & -100 \\ 100 & 101 \end{bmatrix}$$

Clearly $\mathbf{A}^{100}$ can also be obtained by multiplying $\mathbf{A}$ 100 times. However, it is simpler to use Theorem 3.5.

Functions of a square matrix Let $f(\lambda)$ be any function, not necessarily a polynomial. One way to define $f(\mathbf{A})$ is to use Theorem 3.5. Let $h(\lambda)$ be a polynomial of degree $n-1$, where n is the order of $\mathbf{A}$. We solve the coefficients of $h(\lambda)$ by equating $f(\lambda) = h(\lambda)$ on the spectrum of $\mathbf{A}$. Then $f(\mathbf{A})$ is defined as $h(\mathbf{A})$.

Example 3.8 Let

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 0 & -2 \\ 0 & 1 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

Compute $e^{\mathbf{A}_1 t}$. Or, equivalently, if $f(\lambda) = e^{\lambda t}$, what is $f(\mathbf{A}_1)$?

The characteristic polynomial of $\mathbf{A}_1$ is $(\lambda-1)^2(\lambda-2)$. Let $h(\lambda) = \beta_0 + \beta_1\lambda + \beta_2\lambda^2$. Then

$$f(1) = h(1): \quad e^t = \beta_0 + \beta_1 + \beta_2$$
$$f'(1) = h'(1): \quad te^t = \beta_1 + 2\beta_2$$
$$f(2) = h(2): \quad e^{2t} = \beta_0 + 2\beta_1 + 4\beta_2$$

Note that, in the second equation, the differentiation is with respect to λ, not t. Solving these equations yields $\beta_0 = -2te^t + e^{2t}$, $\beta_1 = 3te^t + 2e^t - 2e^{2t}$, and $\beta_2 = e^{2t} - e^t - te^t$. Thus we have

$$e^{\mathbf{A}_1 t} = h(\mathbf{A}_1) = (-2te^t + e^{2t})\mathbf{I} + (3te^t + 2e^t - 2e^{2t})\mathbf{A}_1$$

$$+ (e^{2t} - e^t - te^t)\mathbf{A}_1^2 = \begin{bmatrix} 2e^t - e^{2t} & 0 & 2e^t - 2e^{2t} \\ 0 & e^t & 0 \\ e^{2t} - e^t & 0 & 2e^{2t} - e^t \end{bmatrix}$$

Example 3.9 Let

$$\mathbf{A}_2 = \begin{bmatrix} 0 & 2 & -2 \\ 0 & 1 & 0 \\ 1 & -1 & 3 \end{bmatrix}$$

Compute $e^{\mathbf{A}_2 t}$. The characteristic polynomial of $\mathbf{A}_2$ is $(\lambda-1)^2(\lambda-2)$, which is the same as for $\mathbf{A}_1$. Hence we have the same $h(\lambda)$ as in Example 3.8. Consequently, we have

$$e^{\mathbf{A}_2 t} = h(\mathbf{A}_2) = \begin{bmatrix} 2e^t - e^{2t} & 2te^t & 2e^t - 2e^{2t} \\ 0 & e^t & 0 \\ e^{2t} - e^t & -te^t & 2e^{2t} - e^t \end{bmatrix}$$

EXAMPLE 3.10 Consider the Jordan block of order 4:

$$\hat{\mathbf{A}} = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix} \tag{3.46}$$

Its characteristic polynomial is $(\lambda - \lambda_1)^4$. Although we can select $h(\lambda)$ as $\beta_0 + \beta_1\lambda + \beta_2\lambda + \beta_3\lambda^3$, it is computationally simpler to select $h(\lambda)$ as

$$h(\lambda) = \beta_0 + \beta_1(\lambda - \lambda_1) + \beta_2(\lambda - \lambda_1) + \beta_3(\lambda - \lambda_1)^3$$

This selection is permitted because $h(\lambda)$ has degree $(n - 1) = 3$ and $n = 4$ independent unknowns. The condition $f(\lambda) = h(\lambda)$ on the spectrum of $\hat{\mathbf{A}}$ yields immediately

$$\beta_0 = f(\lambda_1), \quad \beta_1 = f'(\lambda_1), \quad \beta_2 = \frac{f''(\lambda_1)}{2!}, \quad \beta_3 = \frac{f^{(3)}(\lambda_1)}{3!}$$

Thus we have

$$f(\hat{\mathbf{A}}) = f(\lambda_1)\mathbf{I} + \frac{f'(\lambda_1)}{1!}(\hat{\mathbf{A}} - \lambda_1\mathbf{I}) + \frac{f''(\lambda_1)}{2!}(\hat{\mathbf{A}} - \lambda_1\mathbf{I})^2 + \frac{f^{(3)}(\lambda_1)}{3!}(\hat{\mathbf{A}} - \lambda_1\mathbf{I})^3$$

Using the special forms of $(\hat{\mathbf{A}} - \lambda_1\mathbf{I})^k$ as discussed in (3.40), we can readily obtain

$$f(\hat{\mathbf{A}}) = \begin{bmatrix} f(\lambda_1) & f'(\lambda_1)/1! & f''(\lambda_1)/2! & f^{(3)}(\lambda_1)/3! \\ 0 & f(\lambda_1) & f'(\lambda_1)/1! & f''(\lambda_1)/2! \\ 0 & 0 & f(\lambda_1) & f'(\lambda_1)/1! \\ 0 & 0 & 0 & f(\lambda_1) \end{bmatrix} \tag{3.47}$$

If $f(\lambda) = e^{\lambda t}$, then

$$e^{\hat{\mathbf{A}}t} = \begin{bmatrix} e^{\lambda_1 t} & te^{\lambda_1 t} & t^2e^{\lambda_1 t}/2! & t^3e^{\lambda_1 t}/3! \\ 0 & e^{\lambda_1 t} & te^{\lambda_1 t} & t^2e^{\lambda_1 t}/2! \\ 0 & 0 & e^{\lambda_1 t} & te^{\lambda_1 t} \\ 0 & 0 & 0 & e^{\lambda_1 t} \end{bmatrix} \tag{3.48}$$

Because functions of **A** are defined through polynomials of **A**, Equations (3.41) and (3.42) are applicable to functions.

EXAMPLE 3.11 Consider

$$\mathbf{A} = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix}$$

It is block diagonal and contains two Jordan blocks. If $f(\lambda) = e^{\mathbf{A}t}$, then (3.41) and (3.48) imply

$$e^{\mathbf{A}t} = \begin{bmatrix} e^{\lambda_1 t} & te^{\lambda_1 t} & t^2 e^{\lambda_1 t}/2! & 0 & 0 \\ 0 & e^{\lambda_1 t} & te^{\lambda_1 t} & 0 & 0 \\ 0 & 0 & e^{\lambda_1 t} & 0 & 0 \\ 0 & 0 & 0 & e^{\lambda_2 t} & te^{\lambda_2 t} \\ 0 & 0 & 0 & 0 & e^{\lambda_2 t} \end{bmatrix}$$

If $f(\lambda) = (s - \lambda)^{-1}$, then (3.41) and (3.47) imply

$$(s\mathbf{I} - \mathbf{A})^{-1} = \begin{bmatrix} \frac{1}{(s-\lambda_1)} & \frac{1}{(s-\lambda_1)^2} & \frac{1}{(s-\lambda_1)^3} & 0 & 0 \\ 0 & \frac{1}{(s-\lambda_1)} & \frac{1}{(s-\lambda_1)^2} & 0 & 0 \\ 0 & 0 & \frac{1}{(s-\lambda_1)} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{(s-\lambda_2)} & \frac{1}{(s-\lambda_2)^2} \\ 0 & 0 & 0 & 0 & \frac{1}{(s-\lambda_2)} \end{bmatrix} \tag{3.49}$$

Using power series The function of $\mathbf{A}$ was defined using a polynomial of finite degree. We now give an alternative definition by using an infinite power series. Suppose $f(\lambda)$ can be expressed as the power series

$$f(\lambda) = \sum_{i=0}^{\infty} \beta_i \lambda^i$$

with the radius of convergence ρ. If all eigenvalues of $\mathbf{A}$ have magnitudes less than ρ, then $f(\mathbf{A})$ can be defined as

$$f(\mathbf{A}) = \sum_{i=0}^{\infty} \beta_i \mathbf{A}^i \tag{3.50}$$

Instead of proving the equivalence of this definition and the definition based on Theorem 3.5, we use (3.50) to derive (3.47).

Example 3.12 Consider the Jordan-form matrix $\hat{\mathbf{A}}$ in (3.46). Let

$$f(\lambda) = f(\lambda_1) + f'(\lambda_1)(\lambda - \lambda_1) + \frac{f''(\lambda_1)}{2!}(\lambda - \lambda_1)^2 + \cdots$$

then

$$f(\hat{\mathbf{A}}) = f(\lambda_1)\mathbf{I} + f'(\lambda_1)(\hat{\mathbf{A}} - \lambda_1\mathbf{I}) + \cdots + \frac{f^{(n-1)}(\lambda_1)}{(n-1)!}(\hat{\mathbf{A}} - \lambda_1\mathbf{I})^{n-1} + \cdots$$

Because $(\hat{\mathbf{A}} - \lambda_1\mathbf{I})^k = \mathbf{0}$ for $k \geq n = 4$ as discussed in (3.40), the infinite series reduces immediately to (3.47). Thus the two definitions lead to the same function of a matrix.

The most important function of $\mathbf{A}$ is the exponential function $e^{\mathbf{A}t}$. Because the Taylor series

$$e^{\lambda t} = 1 + \lambda t + \frac{\lambda^2 t^2}{2!} + \cdots + \frac{\lambda^n t^n}{n!} + \cdots$$

converges for all finite λ and t, we have

$$e^{\mathbf{A}t} = \mathbf{I} + t\mathbf{A} + \frac{t^2}{2!}\mathbf{A}^2 + \cdots = \sum_{k=0}^{\infty} \frac{1}{k!} t^k \mathbf{A}^k \tag{3.51}$$

This series involves only multiplications and additions and may converge rapidly; therefore it is suitable for computer computation. We list in the following the program in MATLAB that computes (3.51) for $t = 1$:

```
Function E=expm2(A)
E=zeros(size(A));
F=eye(size(A));
k=1;
while norm(E+F-E,1)>0
      E=E+F;
      F=A*F/k;
      k=k+1;
end
```

In the program, E denotes the partial sum and F is the next term to be added to E. The first line defines the function. The next two lines initialize E and F. Let c_k denote the kth term of (3.51) with $t = 1$. Then we have $c_{k+1} = (\mathbf{A}/k)c_k$ for $k = 1, 2, \ldots$. Thus we have F = A * F/k. The computation stops if the 1-norm of E + F − E, denoted by norm(E + F − E, 1), is rounded to 0 in computers. Because the algorithm compares F and E, not F and 0, the algorithm uses norm(E + F − E, 1) instead of norm(F,1). Note that norm(a,1) is the 1-norm discussed in Section 3.2 and will be discussed again in Section 3.9. We see that the series can indeed be programmed easily. To improve the computed result, the techniques of scaling and squaring can be used. In MATLAB, the function `expm2` uses (3.51). The function `expm` or `expm1`, however, uses the so-called Padé approximation. It yields comparable results as `expm2` but requires only about half the computing time. Thus `expm` is preferred to `expm2`. The function `expm3` uses Jordan form, but it will yield an incorrect solution if a matrix is not diagonalizable. If a closed-form solution of $e^{\mathbf{A}t}$ is needed, we must use Theorem 3.5 or Jordan form to compute $e^{\mathbf{A}t}$.

We derive some important properties of $e^{\mathbf{A}t}$ to conclude this section. Using (3.51), we can readily verify the next two equalities

$$e^{\mathbf{0}} = \mathbf{I} \tag{3.52}$$

$$e^{\mathbf{A}(t_1+t_2)} = e^{\mathbf{A}t_1}e^{\mathbf{A}t_2} \tag{3.53}$$

$$[e^{\mathbf{A}t}]^{-1} = e^{-\mathbf{A}t} \tag{3.54}$$

To show (3.54), we set $t_2 = -t_1$. Then (3.53) and (3.52) imply

$$e^{\mathbf{A}t_1}e^{-\mathbf{A}t_1} = e^{\mathbf{A}\cdot 0} = e^{\mathbf{0}} = \mathbf{I}$$

which implies (3.54). Thus the inverse of $e^{\mathbf{A}t}$ can be obtained by simply changing the sign of t. Differentiating term by term of (3.51) yields

$$\begin{aligned}\frac{d}{dt}e^{\mathbf{A}t} &= \sum_{k=1}^{\infty}\frac{1}{(k-1)!}t^{k-1}\mathbf{A}^k \\ &= \mathbf{A}\left(\sum_{k=0}^{\infty}\frac{1}{k!}t^k\mathbf{A}^k\right) = \left(\sum_{k=0}^{\infty}\frac{1}{k!}t^k\mathbf{A}^k\right)\mathbf{A}\end{aligned}$$

Thus we have

$$\frac{d}{dt}e^{\mathbf{A}t} = \mathbf{A}e^{\mathbf{A}t} = e^{\mathbf{A}t}\mathbf{A} \tag{3.55}$$

This is an important equation. We mention that

$$e^{(\mathbf{A}+\mathbf{B})t} \neq e^{\mathbf{A}t}e^{\mathbf{B}t} \tag{3.56}$$

The equality holds only if **A** and **B** commute or **AB** = **BA**. This can be verified by direct substitution of (3.51).

The Laplace transform of a function $f(t)$ is defined as

$$\hat{f}(s) := \mathcal{L}[f(t)] = \int_0^{\infty} f(t)e^{-st}dt$$

It can be shown that

$$\mathcal{L}\left[\frac{t^k}{k!}\right] = s^{-(k+1)}$$

Taking the Laplace transform of (3.51) yields

$$\mathcal{L}[e^{\mathbf{A}t}] = \sum_{k=0}^{\infty}s^{-(k+1)}\mathbf{A}^k = s^{-1}\sum_{k=0}^{\infty}(s^{-1}\mathbf{A})^k$$

Because the infinite series

$$\sum_{k=0}^{\infty}(s^{-1}\lambda)^k = 1 + s^{-1}\lambda + s^{-2}\lambda^2 + \cdots = (1 - s^{-1}\lambda)^{-1}$$

converges for $|s^{-1}\lambda| < 1$, we have

$$\begin{aligned}s^{-1}\sum_{k=0}^{\infty}(s^{-1}\mathbf{A})^k &= s^{-1}\mathbf{I} + s^{-2}\mathbf{A} + s^{-3}\mathbf{A}^2 + \cdots \\ &= s^{-1}(\mathbf{I} - s^{-1}\mathbf{A})^{-1} = \left[s(\mathbf{I} - s^{-1}\mathbf{A})\right]^{-1} = (s\mathbf{I} - \mathbf{A})^{-1}\end{aligned} \tag{3.57}$$

and

$$\mathcal{L}[e^{\mathbf{A}t}] = (s\mathbf{I} - \mathbf{A})^{-1} \tag{3.58}$$

Although in the derivation of (3.57) we require s to be sufficiently large so that all eigenvalues of $s^{-1}\mathbf{A}$ have magnitudes less than 1, Equation (3.58) actually holds for all s except at the eigenvalues of $\mathbf{A}$. Equation (3.58) can also be established from (3.55). Because $\mathcal{L}[df(t)/dt] = s\mathcal{L}[f(t)] - f(0)$, applying the Laplace transform to (3.55) yields

$$s\mathcal{L}[e^{\mathbf{A}t}] - e^{\mathbf{0}} = \mathbf{A}\mathcal{L}[e^{\mathbf{A}t}]$$

or

$$(s\mathbf{I} - \mathbf{A})\mathcal{L}[e^{\mathbf{A}t}] = e^{\mathbf{0}} = \mathbf{I}$$

which implies (3.58).

3.7 Lyapunov Equation

Consider the equation

$$\mathbf{AM} + \mathbf{MB} = \mathbf{C} \tag{3.59}$$

where $\mathbf{A}$ and $\mathbf{B}$ are, respectively, $n \times n$ and $m \times m$ constant matrices. In order for the equation to be meaningful, the matrices $\mathbf{M}$ and $\mathbf{C}$ must be of order $n \times m$. The equation is called the *Lyapunov* equation.

The equation can be written as a set of standard linear algebraic equations. To see this, we assume $n = 3$ and $m = 2$ and write (3.59) explicitly as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{bmatrix} + \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$
$$= \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix}$$

Multiplying them out and then equating the corresponding entries on both sides of the equality, we obtain

$$\begin{bmatrix} a_{11}+b_{11} & a_{12} & a_{13} & b_{21} & 0 & 0 \\ a_{21} & a_{22}+b_{11} & a_{23} & 0 & b_{21} & 0 \\ a_{31} & a_{32} & a_{33}+b_{11} & 0 & 0 & b_{21} \\ b_{12} & 0 & 0 & a_{11}+b_{22} & a_{12} & a_{13} \\ 0 & b_{12} & 0 & a_{21} & a_{22}+b_{22} & a_{23} \\ 0 & 0 & b_{12} & a_{31} & a_{32} & a_{33}+b_{22} \end{bmatrix}$$

$$\times \begin{bmatrix} m_{11} \\ m_{21} \\ m_{31} \\ m_{12} \\ m_{22} \\ m_{32} \end{bmatrix} = \begin{bmatrix} c_{11} \\ c_{21} \\ c_{31} \\ c_{12} \\ c_{22} \\ c_{32} \end{bmatrix} \tag{3.60}$$

This is indeed a standard linear algebraic equation. The matrix on the preceding page is a square matrix of order $n \times m = 3 \times 2 = 6$.

Let us define $\mathcal{A}(\mathbf{M}) := \mathbf{AM} + \mathbf{MB}$. Then the Lyapunov equation can be written as $\mathcal{A}(\mathbf{M}) = \mathbf{C}$. It maps an nm-dimensional linear space into itself. A scalar η is called an eigenvalue of $\mathcal{A}$ if there exists a nonzero $\mathbf{M}$ such that

$$\mathcal{A}(\mathbf{M}) = \eta\mathbf{M}$$

Because $\mathcal{A}$ can be considered as a square matrix of order nm, it has nm eigenvalues η_k, for $k = 1, 2, \ldots, nm$. It turns out

$$\eta_k = \lambda_i + \mu_j \qquad \text{for } i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, m$$

where λ_i, $i = 1, 2, \ldots, n$, and μ_j, $j = 1, 2, \ldots, m$, are, respectively, the eigenvalues of $\mathbf{A}$ and $\mathbf{B}$. In other words, the eigenvalues of $\mathcal{A}$ are all possible sums of the eigenvalues of $\mathbf{A}$ and $\mathbf{B}$.

We show intuitively why this is the case. Let $\mathbf{u}$ be an $n \times 1$ right eigenvector of $\mathbf{A}$ associated with λ_i; that is, $\mathbf{Au} = \lambda_i\mathbf{u}$. Let $\mathbf{v}$ be a $1 \times m$ left eigenvector of $\mathbf{B}$ associated with μ_j; that is, $\mathbf{vB} = \mathbf{v}\mu_j$. Applying $\mathcal{A}$ to the $n \times m$ matrix $\mathbf{uv}$ yields

$$\mathcal{A}(\mathbf{uv}) = \mathbf{Auv} + \mathbf{uvB} = \lambda_i\mathbf{uv} + \mathbf{uv}\mu_j = (\lambda_i + \mu_j)\mathbf{uv}$$

Because both $\mathbf{u}$ and $\mathbf{v}$ are nonzero, so is the matrix $\mathbf{uv}$. Thus $(\lambda_i + \mu_j)$ is an eigenvalue of $\mathcal{A}$.

The determinant of a square matrix is the product of all its eigenvalues. Thus a matrix is nonsingular if and only if it has no zero eigenvalue. If there are no i and j such that $\lambda_i + \mu_j = 0$, then the square matrix in (3.60) is nonsingular and, for every $\mathbf{C}$, there exists a unique $\mathbf{M}$ satisfying the equation. In this case, the Lyapunov equation is said to be nonsingular. If $\lambda_i + \mu_j = 0$ for some i and j, then for a given $\mathbf{C}$, solutions may or may not exist. If $\mathbf{C}$ lies in the range space of $\mathcal{A}$, then solutions exist and are not unique. See Problem 3.32.

The MATLAB function `m=lyap(a,b,-c)` computes the solution of the Lyapunov equation in (3.59).

3.8 Some Useful Formulas

This section discusses some formulas that will be needed later. Let $\mathbf{A}$ and $\mathbf{B}$ be $m \times n$ and $n \times p$ constant matrices. Then we have

$$\rho(\mathbf{AB}) \le \min(\rho(\mathbf{A}), \rho(\mathbf{B})) \tag{3.61}$$

where ρ denotes the rank. This can be argued as follows. Let $\rho(\mathbf{B}) = \alpha$. Then $\mathbf{B}$ has α linearly independent rows. In $\mathbf{AB}$, $\mathbf{A}$ operates on the rows of $\mathbf{B}$. Thus the rows of $\mathbf{AB}$ are linear combinations of the rows of **B**. Thus **AB** has at most α linearly independent rows. In **AB**, **B** operates on the columns of **A**. Thus if **A** has β linearly independent columns, then **AB** has at most β linearly independent columns. This establishes (3.61). Consequently, if $\mathbf{A} = \mathbf{B}_1\mathbf{B}_2\mathbf{B}_3\cdots$, then the rank of **A** is equal to or smaller than the smallest rank of $\mathbf{B}_i$.

Let **A** be $m \times n$ and let **C** and **D** be any $n \times n$ and $m \times m$ nonsingular matrices. Then we have

$$\rho(\mathbf{AC}) = \rho(\mathbf{A}) = \rho(\mathbf{DA}) \tag{3.62}$$

In words, the rank of a matrix will not change after pre- or postmultiplying by a nonsingular matrix. To show (3.62), we define

$$\mathbf{P} := \mathbf{AC} \tag{3.63}$$

Because $\rho(\mathbf{A}) \leq \min(m, n)$ and $\rho(\mathbf{C}) = n$, we have $\rho(\mathbf{A}) \leq \rho(\mathbf{C})$. Thus (3.61) implies

$$\rho(\mathbf{P}) \leq \min(\rho(\mathbf{A}), \rho(\mathbf{C})) \leq \rho(\mathbf{A})$$

Next we write (3.63) as $\mathbf{A} = \mathbf{PC}^{-1}$. Using the same argument, we have $\rho(\mathbf{A}) \leq \rho(\mathbf{P})$. Thus we conclude $\rho(\mathbf{P}) = \rho(\mathbf{A})$. A consequence of (3.62) is that the rank of a matrix will not change by elementary operations. Elementary operations are (1) multiplying a row or a column by a nonzero number, (2) interchanging two rows or two columns, and (3) adding the product of one row (column) and a number to another row (column). These operations are the same as multiplying nonsingular matrices. See Reference [6, p. 542].

Let **A** be $m \times n$ and **B** be $n \times m$. Then we have

$$\det(\mathbf{I}_m + \mathbf{AB}) = \det(\mathbf{I}_n + \mathbf{BA}) \tag{3.64}$$

where $\mathbf{I}_m$ is the unit matrix of order m. To show (3.64), let us define

$$\mathbf{N} = \begin{bmatrix} \mathbf{I}_m & \mathbf{A} \\ \mathbf{0} & \mathbf{I}_n \end{bmatrix} \qquad \mathbf{Q} = \begin{bmatrix} \mathbf{I}_m & \mathbf{0} \\ -\mathbf{B} & \mathbf{I}_n \end{bmatrix} \qquad \mathbf{P} = \begin{bmatrix} \mathbf{I}_m & -\mathbf{A} \\ \mathbf{B} & \mathbf{I}_n \end{bmatrix}$$

We compute

$$\mathbf{NP} = \begin{bmatrix} \mathbf{I}_m + \mathbf{AB} & \mathbf{0} \\ \mathbf{B} & \mathbf{I}_n \end{bmatrix}$$

and

$$\mathbf{QP} = \begin{bmatrix} \mathbf{I}_m & -\mathbf{A} \\ \mathbf{0} & \mathbf{I}_n + \mathbf{BA} \end{bmatrix}$$

Because **N** and **Q** are block triangular, their determinants equal the products of the determinant of their block-diagonal matrices or

$$\det \mathbf{N} = \det \mathbf{I}_m \cdot \det \mathbf{I}_n = 1 = \det \mathbf{Q}$$

Likewise, we have

$$\det(\mathbf{NP}) = \det(\mathbf{I}_m + \mathbf{AB}) \qquad \det(\mathbf{QP}) = \det(\mathbf{I}_n + \mathbf{BA})$$

Because

$$\det(\mathbf{NP}) = \det \mathbf{N} \det \mathbf{P} = \det \mathbf{P}$$

and

$$\det(\mathbf{QP}) = \det \mathbf{Q} \det \mathbf{P} = \det \mathbf{P}$$

we conclude $\det(\mathbf{I}_m + \mathbf{AB}) = \det(\mathbf{I}_n + \mathbf{BA})$.

In N, Q, and P, if $\mathbf{I}_n$, $\mathbf{I}_m$, and $\mathbf{B}$ are replaced, respectively, by $\sqrt{s}\mathbf{I}_n$, $\sqrt{s}\mathbf{I}_m$, and $-\mathbf{B}$, then we can readily obtain

$$s^n \det(s\mathbf{I}_m - \mathbf{AB}) = s^m \det(s\mathbf{I}_n - \mathbf{BA}) \qquad (3.65)$$

which implies, for $n = m$ or for $n \times n$ square matrices $\mathbf{A}$ and $\mathbf{B}$,

$$\det(s\mathbf{I}_n - \mathbf{AB}) = \det(s\mathbf{I}_n - \mathbf{BA}) \qquad (3.66)$$

They are useful formulas.

3.9 Quadratic Form and Positive Definiteness

An $n \times n$ real matrix $\mathbf{M}$ is said to be *symmetric* if its transpose equals itself. The scalar function $\mathbf{x}'\mathbf{M}\mathbf{x}$, where $\mathbf{x}$ is an $n \times 1$ real vector and $\mathbf{M}' = \mathbf{M}$, is called a *quadratic form*. We show that all eigenvalues of symmetric $\mathbf{M}$ are real.

The eigenvalues and eigenvectors of real matrices can be complex as shown in Example 3.6. Therefore we must allow $\mathbf{x}$ to assume complex numbers for the time being and consider the scalar function $\mathbf{x}^*\mathbf{M}\mathbf{x}$, where $\mathbf{x}^*$ is the complex conjugate transpose of $\mathbf{x}$. Taking the complex conjugate transpose of $\mathbf{x}^*\mathbf{M}\mathbf{x}$ yields

$$(\mathbf{x}^*\mathbf{M}\mathbf{x})^* = \mathbf{x}^*\mathbf{M}^*\mathbf{x} = \mathbf{x}^*\mathbf{M}'\mathbf{x} = \mathbf{x}^*\mathbf{M}\mathbf{x}$$

where we have used the fact that the complex conjugate transpose of a real $\mathbf{M}$ reduces to simply the transpose. Thus $\mathbf{x}^*\mathbf{M}\mathbf{x}$ is real for any complex $\mathbf{x}$. This assertion is not true if $\mathbf{M}$ is not symmetric. Let λ be an eigenvalue of $\mathbf{M}$ and $\mathbf{v}$ be its eigenvector; that is, $\mathbf{M}\mathbf{v} = \lambda\mathbf{v}$. Because

$$\mathbf{v}^*\mathbf{M}\mathbf{v} = \mathbf{v}^*\lambda\mathbf{v} = \lambda(\mathbf{v}^*\mathbf{v})$$

and because both $\mathbf{v}^*\mathbf{M}\mathbf{v}$ and $\mathbf{v}^*\mathbf{v}$ are real, the eigenvalue λ must be real. This shows that all eigenvalues of symmetric $\mathbf{M}$ are real. After establishing this fact, we can return our study to exclusively real vector $\mathbf{x}$.

We claim that every symmetric matrix can be diagonalized using a similarity transformation even it has repeated eigenvalue λ. To show this, we show that there is no generalized eigenvector of grade 2 or higher. Suppose $\mathbf{x}$ is a generalized eigenvector of grade 2 or

$$(\mathbf{M} - \lambda\mathbf{I})^2\mathbf{x} = \mathbf{0} \qquad (3.67)$$

$$(\mathbf{M} - \lambda\mathbf{I})\mathbf{x} \neq \mathbf{0} \qquad (3.68)$$

Consider

$$[(\mathbf{M}-\lambda\mathbf{I})\mathbf{x}]'(\mathbf{M}-\lambda\mathbf{I})\mathbf{x} = \mathbf{x}'(\mathbf{M}'-\lambda\mathbf{I}')(\mathbf{M}-\lambda\mathbf{I})\mathbf{x} = \mathbf{x}'(\mathbf{M}-\lambda\mathbf{I})^2\mathbf{x}$$

which is nonzero according to (3.68) but is zero according to (3.67). This is a contradiction. Therefore the Jordan form of **M** has no Jordan block of order 2. Similarly, we can show that the Jordan form of **M** has no Jordan block of order 3 or higher. Thus we conclude that there exists a nonsingular **Q** such that

$$\mathbf{M} = \mathbf{Q}\mathbf{D}\mathbf{Q}^{-1} \tag{3.69}$$

where **D** is a diagonal matrix with real eigenvalues of **M** on the diagonal.

A square matrix **A** is called an *orthogonal matrix* if all columns of **A** are orthonormal. Clearly **A** is nonsingular and we have

$$\mathbf{A}'\mathbf{A} = \mathbf{I} \quad \text{and} \quad \mathbf{A}^{-1} = \mathbf{A}'$$

which imply $\mathbf{A}\mathbf{A}' = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I} = \mathbf{A}'\mathbf{A}$. Thus the inverse of an orthogonal matrix equals its transpose. Consider (3.69). Because $\mathbf{D}' = \mathbf{D}$ and $\mathbf{M}' = \mathbf{M}$, (3.69) equals its own transpose or

$$\mathbf{Q}\mathbf{D}\mathbf{Q}^{-1} = [\mathbf{Q}\mathbf{D}\mathbf{Q}^{-1}]' = [\mathbf{Q}^{-1}]'\mathbf{D}\mathbf{Q}'$$

which implies $\mathbf{Q}^{-1} = \mathbf{Q}'$ and $\mathbf{Q}'\mathbf{Q} = \mathbf{Q}\mathbf{Q}' = \mathbf{I}$. Thus **Q** is an orthogonal matrix; its columns are orthonormalized eigenvectors of **M**. This is summarized as a theorem.

Theorem 3.6

For every real symmetric matrix **M**, there exists an orthogonal matrix **Q** such that

$$\mathbf{M} = \mathbf{Q}\mathbf{D}\mathbf{Q}' \quad \text{or} \quad \mathbf{D} = \mathbf{Q}'\mathbf{M}\mathbf{Q}$$

where **D** is a diagonal matrix with the eigenvalues of **M**, which are all real, on the diagonal.

A symmetric matrix **M** is said to be *positive definite*, denoted by $\mathbf{M} > 0$, if $\mathbf{x}'\mathbf{M}\mathbf{x} > 0$ for every nonzero **x**. It is *positive semidefinite*, denoted by $\mathbf{M} \geq 0$, if $\mathbf{x}'\mathbf{M}\mathbf{x} \geq 0$ for every nonzero **x**. If $\mathbf{M} > 0$, then $\mathbf{x}'\mathbf{M}\mathbf{x} = 0$ if and only if $\mathbf{x} = \mathbf{0}$. If **M** is positive semidefinite, then there exists a nonzero **x** such that $\mathbf{x}'\mathbf{M}\mathbf{x} = 0$. This property will be used repeatedly later.

Theorem 3.7

A symmetric $n \times n$ matrix **M** is positive definite (positive semidefinite) if and only if any one of the following conditions holds.

1. Every eigenvalue of **M** is positive (zero or positive).
2. All the *leading* principal minors of **M** are positive (all the principal minors of **M** are zero or positive).
3. There exists an $n \times n$ nonsingular matrix **N** (an $n \times n$ singular matrix **N** or an $m \times n$ matrix **N** with $m < n$) such that $\mathbf{M} = \mathbf{N}'\mathbf{N}$.

Condition (1) can readily be proved by using Theorem 3.6. Next we consider Conditon (3). If $\mathbf{M} = \mathbf{N}'\mathbf{N}$, then

$$\mathbf{x}'\mathbf{M}\mathbf{x} = \mathbf{x}'\mathbf{N}'\mathbf{N}\mathbf{x} = (\mathbf{N}\mathbf{x})'(\mathbf{N}\mathbf{x}) = ||\mathbf{N}\mathbf{x}||_2^2 \geq 0$$

for any $\mathbf{x}$. If $\mathbf{N}$ is nonsingular, the only $\mathbf{x}$ to make $\mathbf{Nx} = \mathbf{0}$ is $\mathbf{x} = \mathbf{0}$. Thus $\mathbf{M}$ is positive definite. If $\mathbf{N}$ is singular, there exists a nonzero $\mathbf{x}$ to make $\mathbf{Nx} = \mathbf{0}$. Thus $\mathbf{M}$ is positive semidefinite. For a proof of Condition (2), see Reference [10].

We use an example to illustrate the principal minors and leading principal minors. Consider

$$\mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix}$$

Its principal minors are m_{11}, m_{22}, m_{33},

$$\det \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}, \quad \det \begin{bmatrix} m_{11} & m_{13} \\ m_{31} & m_{33} \end{bmatrix}, \quad \det \begin{bmatrix} m_{22} & m_{23} \\ m_{32} & m_{33} \end{bmatrix}$$

and $\det \mathbf{M}$. Thus the principal minors are the determinants of all submatrices of $\mathbf{M}$ whose diagonals coincide with the diagonal of $\mathbf{M}$. The leading principal minors of $\mathbf{M}$ are

$$m_{11}, \quad \det \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix}, \quad \text{and} \quad \det \mathbf{M}$$

Thus the leading principal minors of $\mathbf{M}$ are the determinants of the submatrices of $\mathbf{M}$ obtained by deleting the last k columns and last k rows for $k = 2,\ 1, 0$.

Theorem 3.8

1. An $m \times n$ matrix $\mathbf{H}$, with $m \geq n$, has rank n, if and only if the $n \times n$ matrix $\mathbf{H}'\mathbf{H}$ has rank n or $\det(\mathbf{H}'\mathbf{H}) \neq 0$.

2. An $m \times n$ matrix $\mathbf{H}$, with $m \leq n$, has rank m, if and only if the $m \times m$ matrix $\mathbf{HH}'$ has rank m or $\det(\mathbf{HH}') \neq 0$.

The symmetric matrix $\mathbf{H}'\mathbf{H}$ is always positive semidefinite. It becomes positive definite if $\mathbf{H}'\mathbf{H}$ is nonsingular. We give a proof of this theorem. The argument in the proof will be used to establish the main results in Chapter 6; therefore the proof is spelled out in detail.

Proof: Necessity: The condition $\rho(\mathbf{H}'\mathbf{H}) = n$ implies $\rho(\mathbf{H}) = n$. We show this by contradiction. Suppose $\rho(\mathbf{H}'\mathbf{H}) = n$ but $\rho(\mathbf{H}) < n$. Then there exists a nonzero vector $\mathbf{v}$ such that $\mathbf{Hv} = \mathbf{0}$, which implies $\mathbf{H}'\mathbf{Hv} = \mathbf{0}$. This contradicts $\rho(\mathbf{H}'\mathbf{H}) = n$. Thus $\rho(\mathbf{H}'\mathbf{H}) = n$ implies $\rho(\mathbf{H}) = n$.

Sufficiency: The condition $\rho(\mathbf{H}) = n$ implies $\rho(\mathbf{H}'\mathbf{H}) = n$. Suppose not, or $\rho(\mathbf{H}'\mathbf{H}) < n$; then there exists a nonzero vector $\mathbf{v}$ such that $\mathbf{H}'\mathbf{Hv} = \mathbf{0}$, which implies $\mathbf{v}'\mathbf{H}'\mathbf{Hv} = 0$ or

$$0 = \mathbf{v}'\mathbf{H}'\mathbf{Hv} = (\mathbf{Hv})'(\mathbf{Hv}) = ||\mathbf{Hv}||_2^2$$

Thus we have $\mathbf{Hv} = \mathbf{0}$. This contradicts the hypotheses that $\mathbf{v} \neq 0$ and $\rho(\mathbf{H}) = n$. Thus $\rho(\mathbf{H}) =$ implies $\rho(\mathbf{H}'\mathbf{H}) = n$. This establishes the first part of Theorem 3.8. The second part can be established similarly. Q.E.D.

We discuss the relationship between the eigenvalues of $\mathbf{H}'\mathbf{H}$ and those of $\mathbf{HH}'$. Because both $\mathbf{H}'\mathbf{H}$ and $\mathbf{HH}'$ are symmetric and positive semidefinite, their eigenvalues are real and nonnegative (zero or positive). If $\mathbf{H}$ is $m \times n$, then $\mathbf{H}'\mathbf{H}$ has n eigenvalues and $\mathbf{H}\mathbf{H}'$ has m eigenvalues. Let $\mathbf{A} = \mathbf{H}$ and $\mathbf{B} = \mathbf{H}'$. Then (3.65) becomes

$$\det(s\mathbf{I}_m - \mathbf{H}\mathbf{H}') = s^{m-n}\det(s\mathbf{I}_n - \mathbf{H}'\mathbf{H}) \tag{3.70}$$

This implies that the characteristic polynomials of $\mathbf{H}\mathbf{H}'$ and $\mathbf{H}'\mathbf{H}$ differ only by s^{m-n}. Thus we conclude that $\mathbf{H}\mathbf{H}'$ and $\mathbf{H}'\mathbf{H}$ have the same nonzero eigenvalues but may have different numbers of zero eigenvalues. Furthermore, they have at most $\bar{n} := \min(m, n)$ number of nonzero eigenvalues.

3.10 Singular-Value Decomposition

Let $\mathbf{H}$ be an $m \times n$ real matrix. Define $\mathbf{M} := \mathbf{H}'\mathbf{H}$. Clearly $\mathbf{M}$ is $n \times n$, symmetric, and semidefinite. Thus all eigenvalues of $\mathbf{M}$ are real and nonnegative (zero or positive). Let r be the number of its positive eigenvalues. Then the eigenvalues of $\mathbf{M} = \mathbf{H}'\mathbf{H}$ can be arranged as

$$\lambda_1^2 \geq \lambda_2^2 \geq \cdots \lambda_r^2 > 0 = \lambda_{r+1} = \cdots = \lambda_n$$

Let $\bar{n} := \min(m, n)$. Then the set

$$\lambda_1 \geq \lambda_2 \geq \cdots \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_{\bar{n}}$$

is called the *singular values* of $\mathbf{H}$. The singular values are usually arranged in descending order in magnitude.

EXAMPLE 3.13 Consider the 2×3 matrix

$$\mathbf{H} = \begin{bmatrix} -4 & -1 & 2 \\ 2 & 0.5 & -1 \end{bmatrix}$$

We compute

$$\mathbf{M} = \mathbf{H}'\mathbf{H} = \begin{bmatrix} 20 & 5 & -10 \\ 5 & 1.25 & -2.5 \\ -10 & -2.5 & 5 \end{bmatrix}$$

and compute its characteristic polynomial as

$$\det(\lambda\mathbf{I} - \mathbf{M}) = \lambda^3 - 26.25\lambda^2 = \lambda^2(\lambda - 26.25)$$

Thus the eigenvalues of $\mathbf{H}'\mathbf{H}$ are 26.25, 0, and 0, and the singular values of $\mathbf{H}$ are $\sqrt{26.25} = 5.1235$ and 0. Note that the number of singular values equals min (n, m).

In view of (3.70), we can also compute the singular values of $\mathbf{H}$ from the eigenvalues of $\mathbf{H}\mathbf{H}'$. Indeed, we have

$$\bar{\mathbf{M}} := \mathbf{H}\mathbf{H}' = \begin{bmatrix} 21 & -10.5 \\ -10.5 & 5.25 \end{bmatrix}$$

and

$$\det(\lambda\mathbf{I} - \bar{\mathbf{M}}) = \lambda^2 - 26.25\lambda = \lambda(\lambda - 26.25)$$

Thus the eigenvalues of $\mathbf{H}\mathbf{H}'$ are 26.25 and 0 and the singular values of $\mathbf{H}'$ are 5.1235 and 0. We see that the eigenvalues of $\mathbf{H}'\mathbf{H}$ differ from those of $\mathbf{H}\mathbf{H}'$ only in the number of zero eigenvalues and the singular values of $\mathbf{H}$ equal the singular values of $\mathbf{H}'$.

For $\mathbf{M} = \mathbf{H}'\mathbf{H}$, there exists, following Theorem 3.6, an orthogonal matrix $\mathbf{Q}$ such that

$$\mathbf{Q}'\mathbf{H}'\mathbf{H}\mathbf{Q} = \mathbf{D} =: \mathbf{S}'\mathbf{S} \tag{3.71}$$

where $\mathbf{D}$ is an $n \times n$ diagonal matrix with λ_i^2 on the diagonal. The matrix $\mathbf{S}$ is $m \times n$ with the singular values λ_i on the diagonal. Manipulation on (3.71) will lead eventially to the theorem that follows.

Theorem 3.9 (Singular-value decomposition)

Every $m \times n$ matrix $\mathbf{H}$ can be transformed into the form

$$\mathbf{H} = \mathbf{R}\mathbf{S}\mathbf{Q}'$$

with $\mathbf{R}'\mathbf{R} = \mathbf{R}\mathbf{R}' = \mathbf{I}_m$, $\mathbf{Q}'\mathbf{Q} = \mathbf{Q}\mathbf{Q}' = \mathbf{I}_n$, and $\mathbf{S}$ being $m \times n$ with the singular values of $\mathbf{H}$ on the diagonal.

The columns of $\mathbf{Q}$ are orthonormalized eigenvectors of $\mathbf{H}'\mathbf{H}$ and the columns of $\mathbf{R}$ are orthonormalized eigenvectors of $\mathbf{H}\mathbf{H}'$. Once $\mathbf{R}$, $\mathbf{S}$, and $\mathbf{Q}$ are computed, the rank of $\mathbf{H}$ equals the number of nonzero singular values. If the rank of $\mathbf{H}$ is r, the first r columns of $\mathbf{R}$ are an orthonormal basis of the range space of $\mathbf{H}$. The last $(n - r)$ columns of $\mathbf{Q}$ are an orthonormal basis of the null space of $\mathbf{H}$. Although computing singular-value decomposition is time consuming, it is very reliable and gives a quantitative measure of the rank. Thus it is used in MATLAB to compute the rank, range space, and null space. In MATLAB, the singular values of $\mathbf{H}$ can be obtained by typing `s=svd(H)`. Typing `[R,S,Q]=svd(H)` yields the three matrices in the theorem. Typing `orth(H)` and `null(H)` yields, respectively, orthonormal bases of the range space and null space of $\mathbf{H}$. The function `null` will be used repeatedly in Chapter 7.

EXAMPLE 3.14 Consider the matrix in (3.11). We type

```
a=[0 1 1 2;1 2 3 4;2 0 2 0];
[r,s,q]=svd(a)
```


which yield

$$r = \begin{bmatrix} 0.3782 & -0.3084 & 0.8729 \\ 0.8877 & -0.1468 & -0.4364 \\ 0.2627 & 0.9399 & 0.2182 \end{bmatrix} \qquad s = \begin{bmatrix} 6.1568 & 0 & 0 & 0 \\ 0 & 2.4686 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$q = \begin{bmatrix} 0.2295 & 0.7020 & 0.3434 & -0.5802 \\ 0.3498 & -0.2439 & 0.8384 & 0.3395 \\ 0.5793 & 0.4581 & -0.3434 & 0.5802 \\ 0.6996 & -0.4877 & -0.2475 & -0.4598 \end{bmatrix}$$

Thus the singular values of the matrix $\mathbf{A}$ in (3.11) are 6.1568, 2.4686, and 0. The matrix has two nonzero singular values, thus its rank is 2 and, consequently, its nullity is $4 - \rho(\mathbf{A}) = 2$. The first two columns of r are the orthonormal basis in (3.13) and the last two columns of q are the orthonormal basis in (3.14).

3.11 Norms of Matrices

The concept of norms for vectors can be extended to matrices. This concept is needed in Chapter 5. Let **A** be an $m \times n$ matrix. The norm of **A** can be defined as

$$||\mathbf{A}|| = \sup_{\mathbf{x}\neq 0} \frac{||\mathbf{Ax}||}{||\mathbf{x}||} = \sup_{||\mathbf{x}||=1} ||\mathbf{Ax}|| \qquad (3.72)$$

where sup stands for supremum or the least upper bound. This norm is defined through the norm of **x** and is therefore called an *induced norm*. For different $||\mathbf{x}||$, we have different $||\mathbf{A}||$. For example, if the 1-norm $||\mathbf{x}||_1$ is used, then

$$||\mathbf{A}||_1 = \max_j \left(\sum_{i=1}^{m} |a_{ij}| \right) = \text{largest column absolute sum}$$

where a_{ij} is the ijth element of **A**. If the Euclidean norm $||\mathbf{x}||_2$ is used, then

$$\begin{aligned} ||\mathbf{A}||_2 &= \text{largest singular value of } \mathbf{A} \\ &= (\text{largest eigenvalue of } \mathbf{A}'\mathbf{A})^{1/2} \end{aligned}$$

If the infinite-norm $||\mathbf{x}||_\infty$ is used, then

$$||\mathbf{A}||_\infty = \max_i \left(\sum_{j=1}^{n} |a_{ij}| \right) = \text{largest row absolute sum}$$

These norms are all different for the same **A**. For example, if

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

then $||\mathbf{A}||_1 = 3 + |-1| = 4$, $||\mathbf{A}||_2 = 3.7$, and $||\mathbf{A}||_\infty = 3 + 2 = 5$, as shown in Fig. 3.3. The MATLAB functions `norm(a,1)`, `norm(a,2)=norm(a)`, and `norm(a,inf)` compute the three norms.

The norm of matrices has the following properties:

$$||\mathbf{Ax}|| \le ||\mathbf{A}||||\mathbf{x}||$$

$$||\mathbf{A}+\mathbf{B}|| \le ||\mathbf{A}|| + ||\mathbf{B}||$$

$$||\mathbf{AB}|| \le ||\mathbf{A}||||\mathbf{B}||$$

PROBLEMS

The reader should try first to solve all problems involving numerical numbers by hand and then verify the results using MATLAB or any software.

3.1 Consider Fig. 3.1. What is the representation of the vector **x** with respect to the basis $\{\mathbf{q}_1, \mathbf{i}_2\}$? What is the representation of $\mathbf{q}_1$ with respect to $\{\mathbf{i}_2, \mathbf{q}_2\}$?

3.2 What are the 1-norm, 2-norm, and infinite-norm of the vectors

Figure 3.3 Di

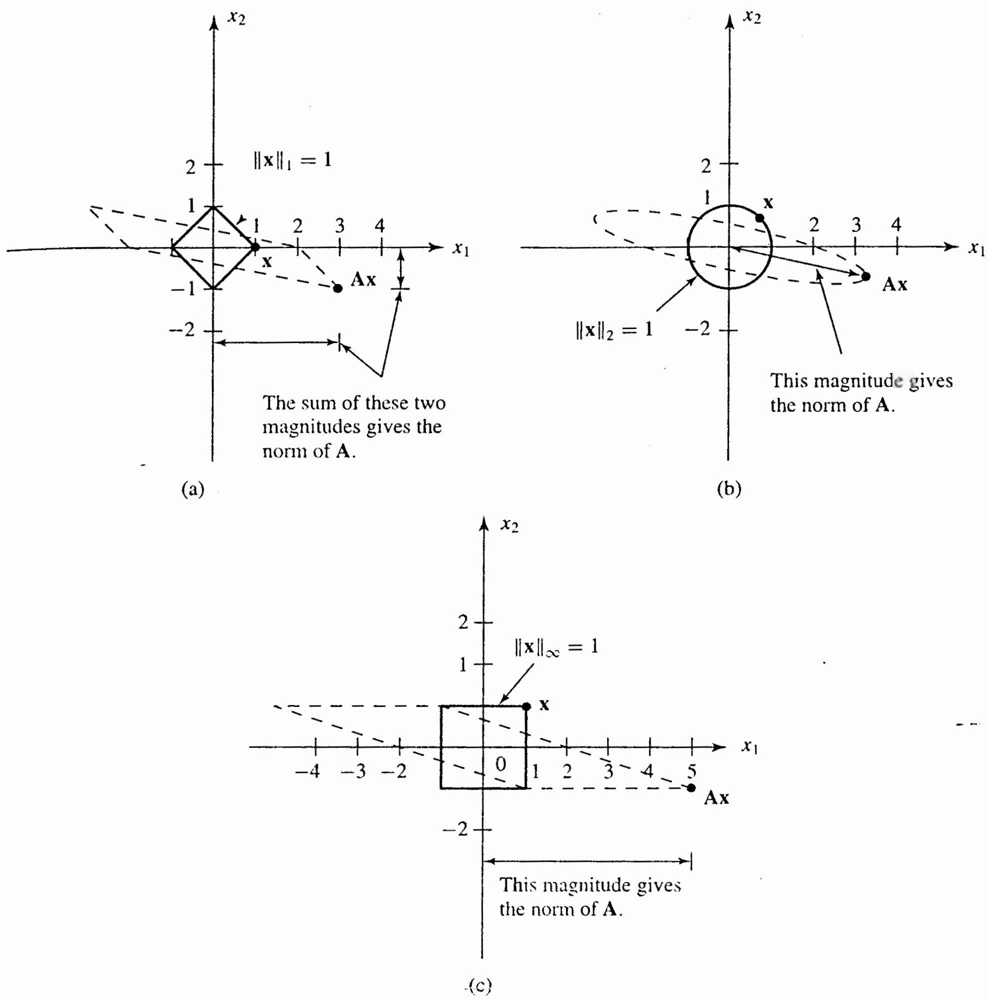

Figure 3.3 Different norms of **A**.

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} \qquad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

3.3 Find two orthonormal vectors that span the same space as the two vectors in Problem 3.2.

3.4 Consider an $n \times m$ matrix $\mathbf{A}$ with $n \geq m$. If all columns of $\mathbf{A}$ are orthonormal, then $\mathbf{A}'\mathbf{A} = \mathbf{I}_m$. What can you say about $\mathbf{A}\mathbf{A}'$?

3.5 Find the ranks and nullities of the following matrices:

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{A}_2 = \begin{bmatrix} 4 & 1 & -1 \\ 3 & 2 & 0 \\ 1 & 1 & 0 \end{bmatrix} \quad \mathbf{A}_3 = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

3.6 Find bases of the range spaces and null spaces of the matrices in Problem 3.5.

3.7 Consider the linear algebraic equation

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

It has three equations and two unknowns. Does a solution $\mathbf{x}$ exist in the equation? Is the solution unique? Does a solution exist if $\mathbf{y} = [1 \;\; 1 \;\; 1]'$?

3.8 Find the general solution of

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$

How many parameters do you have?

3.9 Find the solution in Example 3.3 that has the smallest Euclidean norm.

3.10 Find the solution in Problem 3.8 that has the smallest Euclidean norm.

3.11 Consider the equation

$$\mathbf{x}[n] = \mathbf{A}^n x[0] + \mathbf{A}^{n-1}\mathbf{b}u[0] + \mathbf{A}^{n-2}\mathbf{b}u[1] + \cdots + \mathbf{A}\mathbf{b}u[n-2] + \mathbf{b}u[n-1]$$

where $\mathbf{A}$ is an $n \times n$ matrix and $\mathbf{b}$ is an $n \times 1$ column vector. Under what conditions on $\mathbf{A}$ and $\mathbf{b}$ will there exist $u[0], u[1], \ldots, u[n-1]$ to meet the equation for any $\mathbf{x}[n]$ and $\mathbf{x}[0]$? [*Hint:* Write the equation in the form

$$\mathbf{x}[n] - \mathbf{A}^n\mathbf{x}[0] = [\mathbf{b} \;\; \mathbf{A}\mathbf{b} \;\; \cdots \;\; \mathbf{A}^{n-1}\mathbf{b}] \begin{bmatrix} u[n-1] \\ u[n-2] \\ \vdots \\ u[0] \end{bmatrix}]$$

3.12 Given

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \quad \bar{\mathbf{b}} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix}$$

what are the representations of $\mathbf{A}$ with respect to the basis $\{\mathbf{b}, \mathbf{A}\mathbf{b}, \mathbf{A}^2\mathbf{b}, \mathbf{A}^3\mathbf{b}\}$ and the basis $\{\bar{\mathbf{b}}, \mathbf{A}\bar{\mathbf{b}}, \mathbf{A}^2\bar{\mathbf{b}}, \mathbf{A}^3\bar{\mathbf{b}}\}$, respectively? (Note that the representations are the same!)

3.13 Find Jordan-form representations of the following matrices:

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 4 & 10 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \qquad \mathbf{A}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -4 & -3 \end{bmatrix}$$

$$\mathbf{A}_3 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad \mathbf{A}_4 = \begin{bmatrix} 0 & 4 & 3 \\ 0 & 20 & 16 \\ 0 & -25 & -20 \end{bmatrix}$$

Note that all except $\mathbf{A}_4$ can be diagonalized.

3.14 Consider the companion-form matrix

$$\mathbf{A} = \begin{bmatrix} -\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Show that its characteristic polynomial is given by

$$\Delta(\lambda) = \lambda^4 + \alpha_1\lambda^3 + \alpha_2\lambda^2 + \alpha_3\lambda + \alpha_4$$

Show also that if λ_i is an eigenvalue of $\mathbf{A}$ or a solution of $\Delta(\lambda) = 0$, then $[\lambda_i^3 \;\; \lambda_i^2 \;\; \lambda_i \;\; 1]'$ is an eigenvector of $\mathbf{A}$ associated with λ_i.

3.15 Show that the *Vandermonde* determinant

$$\begin{bmatrix} \lambda_1^3 & \lambda_2^3 & \lambda_3^3 & \lambda_4^3 \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \lambda_4^2 \\ \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

equals $\prod_{1 \le i < j \le 4}(\lambda_j - \lambda_i)$. Thus we conclude that the matrix is nonsingular or, equivalently, the eigenvectors are linearly independent if all eigenvalues are distinct.

3.16 Show that the companion-form matrix in Problem 3.14 is nonsingular if and only if $\alpha_4 \neq 0$. Under this assumption, show that its inverse equals

$$\mathbf{A}^{-1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1/\alpha_4 & -\alpha_1/\alpha_4 & -\alpha_2/\alpha_4 & -\alpha_3/\alpha_4 \end{bmatrix}$$

3.17 Consider

$$\mathbf{A} = \begin{bmatrix} \lambda & \lambda T & \lambda T^2/2 \\ 0 & \lambda & \lambda T \\ 0 & 0 & \lambda \end{bmatrix}$$

with $\lambda \neq 0$ and $T > 0$. Show that $[0 \;\; 0 \;\; 1]'$ is a generalized eigenvector of grade 3 and the three columns of

$$\mathbf{Q} = \begin{bmatrix} \lambda^2 T^2 & \lambda T^2 & 0 \\ 0 & \lambda T & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

constitute a chain of generalized eigenvectors of length 3. Verify

$$\mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$$

3.18 Find the characteristic polynomials and the minimal polynomials of the following matrices:

$$\begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{bmatrix} \quad \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix}$$

$$\begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix} \quad \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix}$$

3.19 Show that if λ is an eigenvalue of $\mathbf{A}$ with eigenvector $\mathbf{x}$, then $f(\lambda)$ is an eigenvalue of $f(\mathbf{A})$ with the same eigenvector $\mathbf{x}$.

3.20 Show that an $n \times n$ matrix has the property $\mathbf{A}^k = \mathbf{0}$ for $k \geq m$ if and only if $\mathbf{A}$ has eigenvalues 0 with multiplicity n and index m or less. Such a matrix is called a *nilpotent* matrix.

3.21 Given

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

find $\mathbf{A}^{10}$, $\mathbf{A}^{103}$, and $e^{\mathbf{A}t}$.

3.22 Use two different methods to compute $e^{\mathbf{A}t}$ for $\mathbf{A}_1$ and $\mathbf{A}_4$ in Problem 3.13.

3.23 Show that functions of the same matrix commute; that is,

$$f(\mathbf{A})g(\mathbf{A}) = g(\mathbf{A})f(\mathbf{A})$$

Consequently we have $\mathbf{A}e^{\mathbf{A}t} = e^{\mathbf{A}t}\mathbf{A}$.

3.24 Let

$$\mathbf{C} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

Find a matrix $\mathbf{B}$ such that $e^{\mathbf{B}} = \mathbf{C}$. Show that if $\lambda_i = 0$ for some i, then $\mathbf{B}$ does not exist. Let

$$\mathbf{C} = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{bmatrix}$$

Find a $\mathbf{B}$ such that $e^{\mathbf{B}} = \mathbf{C}$. Is it true that, for any nonsingular $\mathbf{C}$, there exists a matrix $\mathbf{B}$ such that $e^{\mathbf{B}} = \mathbf{C}$?

3.25 Let

$$(s\mathbf{I} - \mathbf{A})^{-1} = \frac{1}{\Delta(s)} \text{Adj}\,(s\mathbf{I} - \mathbf{A})$$

and let $m(s)$ be the monic greatest common divisor of all entries of Adj $(s\mathbf{I} - \mathbf{A})$. Verify for the matrix $\mathbf{A}_3$ in Problem 3.13 that the minimal polynominal of $\mathbf{A}$ equals $\Delta(s)/m(s)$.

3.26 Define

$$(s\mathbf{I} - \mathbf{A})^{-1} := \frac{1}{\Delta(s)} \left[\mathbf{R}_0 s^{n-1} + \mathbf{R}_1 s^{n-2} + \cdots + \mathbf{R}_{n-2} s + \mathbf{R}_{n-1}\right]$$

where

$$\Delta(s) := \det(s\mathbf{I} - \mathbf{A}) := s^n + \alpha_1 s^{n-1} + \alpha_2 s^{n-2} + \cdots + \alpha_n$$

and $\mathbf{R}_i$ are constant matrices. This definition is valid because the degree in s of the adjoint of $(s\mathbf{I} - \mathbf{A})$ is at most $n - 1$. Verify

$$\alpha_1 = -\frac{\text{tr}(\mathbf{A}\mathbf{R}_0)}{1} \qquad \mathbf{R}_0 = \mathbf{I}$$

$$\alpha_2 = -\frac{\text{tr}(\mathbf{A}\mathbf{R}_1)}{2} \qquad \mathbf{R}_1 = \mathbf{A}\mathbf{R}_0 + \alpha_1\mathbf{I} = \mathbf{A} + \alpha_1\mathbf{I}$$

$$\alpha_3 = -\frac{\text{tr}(\mathbf{A}\mathbf{R}_2)}{3} \qquad \mathbf{R}_2 = \mathbf{A}\mathbf{R}_1 + \alpha_2\mathbf{I} = \mathbf{A}^2 + \alpha_1\mathbf{A} + \alpha_2\mathbf{I}$$

$$\vdots$$

$$\alpha_{n-1} = -\frac{\text{tr}(\mathbf{A}\mathbf{R}_{n-2})}{n-1} \qquad \mathbf{R}_{n-1} = \mathbf{A}\mathbf{R}_{n-2} + \alpha_{n-1}\mathbf{I} = \mathbf{A}^{n-1} + \alpha_1\mathbf{A}^{n-2} + \cdots + \alpha_{n-2}\mathbf{A} + \alpha_{n-1}\mathbf{I}$$

$$\alpha_n = -\frac{\text{tr}(\mathbf{A}\mathbf{R}_{n-1})}{n} \qquad \mathbf{0} = \mathbf{A}\mathbf{R}_{n-1} + \alpha_n\mathbf{I}$$

where tr stands for the *trace* of a matrix and is defined as the sum of all its diagonal entries. This process of computing α_i and $\mathbf{R}_i$ is called the *Leverrier algorithm*.

3.27 Use Problem 3.26 to prove the Cayley–Hamilton theorem.

3.28 Use Problem 3.26 to show

$$(s\mathbf{I} - \mathbf{A})^{-1} = \frac{1}{\Delta(s)} \left[\mathbf{A}^{n-1} + (s + \alpha_1)\mathbf{A}^{n-2} + (s^2 + \alpha_1 s + \alpha_2)\mathbf{A}^{n-3}\right.$$

$$+\cdots+(s^{n-1}+\alpha_1 s^{n-2}+\cdots+\alpha_{n-1})\mathbf{I}]$$

3.29 Let all eigenvalues of $\mathbf{A}$ be distinct and let $\mathbf{q}_i$ be a right eigenvector of $\mathbf{A}$ associated with λ_i; that is, $\mathbf{A}\mathbf{q}_i = \lambda_i \mathbf{q}_i$. Define $\mathbf{Q} = [\mathbf{q}_1\ \ \mathbf{q}_2\ \ \cdots\ \ \mathbf{q}_n]$ and define

$$\mathbf{P} := \mathbf{Q}^{-1} =: \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \vdots \\ \mathbf{p}_n \end{bmatrix}$$

where $\mathbf{p}_i$ is the ith row of $\mathbf{P}$. Show that $\mathbf{p}_i$ is a left eigenvector of $\mathbf{A}$ associated with λ_i; that is, $\mathbf{p}_i\mathbf{A} = \lambda_i \mathbf{p}_i$.

3.30 Show that if all eigenvalues of $\mathbf{A}$ are distinct, then $(s\mathbf{I}-\mathbf{A})^{-1}$ can be expressed as

$$(s\mathbf{I}-\mathbf{A})^{-1} = \sum \frac{1}{s-\lambda_i}\mathbf{q}_i\mathbf{p}_i$$

where $\mathbf{q}_i$ and $\mathbf{p}_i$ are right and left eigenvectors of $\mathbf{A}$ associated with λ_i.

3.31 Find the $\mathbf{M}$ to meet the Lyapunov equation in (3.59) with

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

What are the eigenvalues of the Lyapunov equation? Is the Lyapunov equation singular? Is the solution unique?

3.32 Repeat Problem 3.31 for

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ -1 & -2 \end{bmatrix} \qquad \mathbf{B} = 1 \qquad \mathbf{C}_1 = \begin{bmatrix} 3 \\ 3 \end{bmatrix} \qquad \mathbf{C}_2 = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$$

with two different $\mathbf{C}$.

3.33 Check to see if the following matrices are positive definite or semidefinite:

$$\begin{bmatrix} 2 & 3 & 2 \\ 3 & 1 & 0 \\ 2 & 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} a_1a_1 & a_1a_2 & a_1a_3 \\ a_2a_1 & a_2a_2 & a_2a_3 \\ a_3a_1 & a_3a_2 & a_3a_3 \end{bmatrix}$$

3.34 Compute the singular values of the following matrices:

$$\begin{bmatrix} -1 & 0 & 1 \\ 2 & -1 & 0 \end{bmatrix} \qquad \begin{bmatrix} -1 & 2 \\ 2 & 4 \end{bmatrix}$$

3.35 If $\mathbf{A}$ is symmetric, what is the relationship between its eigenvalues and singular values?

3.36 Show

$$\det\left(\mathbf{I}_n + \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} [b_1\ b_2\ \cdots\ b_n]\right) = 1 + \sum_{m=1}^{n} a_m b_m$$

3.37 Show (3.65).

3.38 Consider $\mathbf{Ax} = \mathbf{y}$, where $\mathbf{A}$ is $m \times n$ and has rank m. Is $(\mathbf{A}'\mathbf{A})^{-1}\mathbf{A}'\mathbf{y}$ a solution? If not, under what condition will it be a solution? Is $\mathbf{A}'(\mathbf{AA}')^{-1}\mathbf{y}$ a solution?

Chapter

4

State-Space Solutions and Realizations

4.1 Introduction

We showed in Chapter 2 that linear systems can be described by convolutions and, if lumped, by state-space equations. This chapter discusses how to find their solutions. First we discuss briefly how to compute solutions of the input-output description. There is no simple analytical way of computing the convolution

$$y(t) = \int_{\tau=t_0}^{t} g(t, \tau)u(\tau)\, d\tau$$

The easiest way is to compute it numerically on a digital computer. Before doing so, the equation must be discretized. One way is to discretize it as

$$y(k\Delta) = \sum_{m=k_0}^{k} g(k\Delta, m\Delta)u(m\Delta)\Delta \tag{4.1}$$

where Δ is called the integration step size. This is basically the discrete convolution discussed in (2.34). This discretization is the easiest but yields the least accurate result for the same integration step size. For other integration methods, see, for example, Reference [17].

For the linear time-invariant (LTI) case, we can also use $\hat{y}(s) = \hat{g}(s)\hat{u}(s)$ to compute the solution. If a system is distributed, $\hat{g}(s)$ will not be a rational function of s. Except for some special cases, it is simpler to compute the solution directly in the time domain as in (4.1). If the system is lumped, $\hat{g}(s)$ will be a rational function of s. In this case, if the Laplace transform of $u(t)$ is also a rational function of s, then the solution can be obtained by taking the inverse Laplace transform of $\hat{g}(s)\hat{u}(s)$. This method requires computing poles, carrying out partial fraction expansion, and then using a Laplace transform table. These can be carried out using the MATLAB functions `roots` and `residue`. However, when there are repeated poles, the computation may become very sensitive to small changes in the data, including roundoff errors; therefore computing solutions using the Laplace transform is not a viable method on digital computers. A better method is to transform transfer functions into state-space equations and then compute the solutions. This chapter discusses solutions of state equations, how to transform transfer functions into state equations, and other related topics. We discuss first the time-invariant case and then the time-varying case.

4.2 Solution of LTI State Equations

Consider the linear time-invariant (LTI) state-space equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \tag{4.2}$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \tag{4.3}$$

where **A**, **B**, **C**, and **D** are, respectively, $n \times n$, $n \times p$, $q \times n$, and $q \times p$ constant matrices. The problem is to find the solution excited by the initial state $\mathbf{x}(0)$ and the input $\mathbf{u}(t)$. The solution hinges on the exponential function of **A** studied in Section 3.6. In particular, we need the property in (3.55) or

$$\frac{d}{dt}e^{\mathbf{A}t} = \mathbf{A}e^{\mathbf{A}t} = e^{\mathbf{A}t}\mathbf{A}$$

to develop the solution. Premultiplying $e^{-\mathbf{A}t}$ on both sides of (4.2) yields

$$e^{-\mathbf{A}t}\dot{\mathbf{x}}(t) - e^{-\mathbf{A}t}\mathbf{A}\mathbf{x}(t) = e^{-\mathbf{A}t}\mathbf{B}\mathbf{u}(t)$$

which implies

$$\frac{d}{dt}\left(e^{-\mathbf{A}t}\mathbf{x}(t)\right) = e^{-\mathbf{A}t}\mathbf{B}\mathbf{u}(t)$$

Its integration from 0 to t yields

$$e^{-\mathbf{A}\tau}\mathbf{x}(\tau)\Big|_{\tau=0}^{t} = \int_0^t e^{-\mathbf{A}\tau}\mathbf{B}\mathbf{u}(\tau)\,d\tau$$

Thus we have

$$e^{-\mathbf{A}t}\mathbf{x}(t) - e^{\mathbf{0}}\mathbf{x}(0) = \int_0^t e^{-\mathbf{A}\tau}\mathbf{B}\mathbf{u}(\tau)\,d\tau \tag{4.4}$$

Because the inverse of $e^{-\mathbf{A}t}$ is $e^{\mathbf{A}t}$ and $e^{\mathbf{0}} = \mathbf{I}$ as discussed in (3.54) and (3.52), (4.4) implies

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\,d\tau \tag{4.5}$$

This is the solution of (4.2).

It is instructive to verify that (4.5) is the solution of (4.2). To verify this, we must show that (4.5) satisfies (4.2) and the initial condition $\mathbf{x}(t) = \mathbf{x}(0)$ at $t = 0$. Indeed, at $t = 0$, (4.5) reduces to

$$\mathbf{x}(0) = e^{\mathbf{A}\cdot 0}\mathbf{x}(0) = e^{0}\mathbf{x}(0) = \mathbf{I}\mathbf{x}(0) = \mathbf{x}(0)$$

Thus (4.5) satisfies the initial condition. We need the equation

$$\frac{\partial}{\partial t}\int_{t_0}^{t} f(t,\tau)\,d\tau = \int_{t_0}^{t}\left(\frac{\partial}{\partial t} f(t,\tau)\right)d\tau + f(t,\tau)|_{\tau=t} \tag{4.6}$$

to show that (4.5) satisfies (4.2). Differentiating (4.5) and using (4.6), we obtain

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \frac{d}{dt}\left[e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\,d\tau\right] \\ &= \mathbf{A}e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t \mathbf{A}e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\,d\tau + e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\big|_{\tau=t} \\ &= \mathbf{A}\left(e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\,d\tau\right) + e^{\mathbf{A}\cdot 0}\mathbf{B}\mathbf{u}(t)\end{aligned}$$

which becomes, after substituting (4.5),

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$$

Thus (4.5) meets (4.2) and the initial condition $\mathbf{x}(0)$ and is the solution of (4.2).

Substituting (4.5) into (4.3) yields the solution of (4.3) as

$$\mathbf{y}(t) = \mathbf{C}e^{\mathbf{A}t}\mathbf{x}(0) + \mathbf{C}\int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}\mathbf{u}(\tau)\,d\tau + \mathbf{D}\mathbf{u}(t) \tag{4.7}$$

This solution and (4.5) are computed directly in the time domain. We can also compute the solutions by using the Laplace transform. Applying the Laplace transform to (4.2) and (4.3) yields, as derived in (2.14) and (2.15),

$$\begin{aligned}\hat{\mathbf{x}}(s) &= (s\mathbf{I} - \mathbf{A})^{-1}[\mathbf{x}(0) + \mathbf{B}\hat{\mathbf{u}}(s)] \\ \hat{\mathbf{y}}(s) &= \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}[\mathbf{x}(0) + \mathbf{B}\hat{\mathbf{u}}(s)] + \mathbf{D}\hat{\mathbf{u}}(s)\end{aligned}$$

Once $\hat{\mathbf{x}}(s)$ and $\hat{\mathbf{y}}(s)$ are computed algebraically, their inverse Laplace transforms yield the time-domain solutions.

We now give some remarks concerning the computation of $e^{\mathbf{A}t}$. We discussed in Section 3.6 three methods of computing functions of a matrix. They can all be used to compute $e^{\mathbf{A}t}$:

1. Using Theorem 3.5: First, compute the eigenvalues of $\mathbf{A}$; next, find a polynomial $h(\lambda)$ of degree $n-1$ that equals $e^{\lambda t}$ on the spectrum of $\mathbf{A}$; then $e^{\mathbf{A}t} = h(\mathbf{A})$.
2. Using Jordan form of $\mathbf{A}$: Let $\mathbf{A} = \mathbf{Q}\hat{\mathbf{A}}\mathbf{Q}^{-1}$; then $e^{\mathbf{A}t} = \mathbf{Q}e^{\hat{\mathbf{A}}t}\mathbf{Q}^{-1}$, where $\hat{\mathbf{A}}$ is in Jordan form and $e^{\hat{\mathbf{A}}t}$ can readily be obtained by using (3.48).
3. Using the infinite power series in (3.51): Although the series will not, in general, yield a closed-form solution, it is suitable for computer computation, as discussed following (3.51).

In addition, we can use (3.58) to compute $e^{\mathbf{A}t}$, that is,

$$e^{\mathbf{A}t} = \mathcal{L}^{-1}(s\mathbf{I} - \mathbf{A})^{-1} \tag{4.8}$$

The inverse of $(s\mathbf{I}-\mathbf{A})$ is a function of $\mathbf{A}$; therefore, again, we have many methods to compute it:

1. Taking the inverse of $(s\mathbf{I}-\mathbf{A})$.
2. Using Theorem 3.5.
3. Using $(s\mathbf{I}-\mathbf{A})^{-1} = \mathbf{Q}(s\mathbf{I}-\hat{\mathbf{A}})^{-1}\mathbf{Q}^{-1}$ and (3.49).
4. Using the infinite power series in (3.57).
5. Using the Leverrier algorithm discussed in Problem 3.26.

Example 4.1 We use Methods 1 and 2 to compute $(s\mathbf{I}-\mathbf{A})^{-1}$, where

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

Method 1: We use (3.20) to compute

$$(s\mathbf{I}-\mathbf{A})^{-1} = \begin{bmatrix} s & 1 \\ -1 & s+2 \end{bmatrix}^{-1} = \frac{1}{s^2+2s+1}\begin{bmatrix} s+2 & -1 \\ 1 & s \end{bmatrix}$$
$$= \begin{bmatrix} (s+2)/(s+1)^2 & -1/(s+1)^2 \\ 1/(s+1)^2 & s/(s+)^2 \end{bmatrix}$$

Method 2: The eigenvalues of $\mathbf{A}$ are $-1, -1$. Let $h(\lambda) = \beta_0 + \beta_1\lambda$. If $h(\lambda)$ equals $f(\lambda) := (s-\lambda)^{-1}$ on the spectrum of $\mathbf{A}$, then

$$f(-1) = h(-1): \qquad (s+1)^{-1} = \beta_0 - \beta_1$$
$$f'(-1) = h'(-1): \qquad (s+1)^{-2} = \beta_1$$

Thus we have

$$h(\lambda) = [(s+1)^{-1} + (s+1)^{-2}] + (s+1)^{-2}\lambda$$

and

$$(s\mathbf{I}-\mathbf{A})^{-1} = h(\mathbf{A}) = [(s+1)^{-1} + (s+1)^{-2}]\mathbf{I} + (s+1)^{-2}\mathbf{A}$$
$$= \begin{bmatrix} (s+2)/(s+1)^2 & -1/(s+1)^2 \\ 1/(s+1)^2 & s/(s+)^2 \end{bmatrix}$$

Example 4.2 Consider the equation

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 0 & -1 \\ 1 & -2 \end{bmatrix}\mathbf{x}(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix}u(t)$$

Its solution is

$$\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0) + \int_0^t e^{\mathbf{A}(t-\tau)}\mathbf{B}u(\tau)\,d\tau$$

The matrix function $e^{\mathbf{A}t}$ is the inverse Laplace transform of $(s\mathbf{I}-\mathbf{A})^{-1}$, which was computed in the preceding example. Thus we have

$$e^{\mathbf{A}t} = \mathcal{L}^{-1}\begin{bmatrix} \dfrac{s+2}{(s+1)^2} & \dfrac{-1}{(s+1)^2} \\ \dfrac{1}{(s+1)^2} & \dfrac{s}{(s+)^2} \end{bmatrix} = \begin{bmatrix} (1+t)e^{-t} & -te^{-t} \\ te^{-t} & (1-t)e^{-t} \end{bmatrix}$$

and

$$\mathbf{x}(t) = \begin{bmatrix} (1+t)e^{-t} & -te^{-t} \\ te^{-t} & (1-t)e^{-t} \end{bmatrix}\mathbf{x}(0) + \begin{bmatrix} -\int_0^t (t-\tau)e^{-(t-\tau)}u(\tau)\,d\tau \\ \int_0^t [1-(t-\tau)]e^{-(t-\tau)}u(\tau)\,d\tau \end{bmatrix}$$

We discuss a general property of the zero-input response $e^{\mathbf{A}t}\mathbf{x}(0)$. Consider the second matrix in (3.39). Then we have

$$e^{\mathbf{A}t} = \mathbf{Q}\begin{bmatrix} e^{\lambda_1 t} & te^{\lambda_1 t} & t^2e^{\lambda_1 t}/2 & 0 & 0 \\ 0 & e^{\lambda_1 t} & te^{\lambda_1 t} & 0 & 0 \\ 0 & 0 & e^{\lambda_1 t} & 0 & 0 \\ 0 & 0 & 0 & e^{\lambda_1 t} & 0 \\ 0 & 0 & 0 & 0 & e^{\lambda_2 t} \end{bmatrix}\mathbf{Q}^{-1}$$

Every entry of $e^{\mathbf{A}t}$ and, consequently, of the zero-input response is a linear combination of terms $\{e^{\lambda_1 t},\ te^{\lambda_1 t},\ t^2e^{\lambda_1 t},\ e^{\lambda_2 t}\}$. These terms are dictated by the eigenvalues and their indices. In general, if $\mathbf{A}$ has eigenvalue λ_1 with index $\bar{n}_1$, then every entry of $e^{\mathbf{A}t}$ is a linear combination of

$$e^{\lambda_1 t} \quad te^{\lambda_1 t} \quad \cdots \quad t^{\bar{n}_1 - 1}e^{\lambda_1 t}$$

Every such term is *analytic* in the sense that it is infinitely differentiable and can be expanded in a Taylor series at every t. This is a nice property and will be used in Chapter 6.

If every eigenvalue, simple or repeated, of $\mathbf{A}$ has a negative real part, then every zero-input response will approach zero as $t \to \infty$. If $\mathbf{A}$ has an eigenvalue, simple or repeated, with a positive real part, then most zero-input responses will grow unbounded as $t \to \infty$. If $\mathbf{A}$ has some eigenvalues with zero real part and all with index 1 and the remaining eigenvalues all have negative real parts, then no zero-input response will grow unbounded. However, if the index is 2 or higher, then some zero-input response may become unbounded. For example, if $\mathbf{A}$ has eigenvalue 0 with index 2, then $e^{\mathbf{A}t}$ contains the terms $\{1,\ t\}$. If a zero-input response contains the term t, then it will grow unbounded.

4.2.1 Discretization

Consider the continuous-time state equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \tag{4.9}$$

$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \tag{4.10}$$

If the set of equations is to be computed on a digital computer, it must be discretized. Because

$$\dot{\mathbf{x}}(t) = \lim_{T\to 0}\frac{\mathbf{x}(t+T) - \mathbf{x}(t)}{T}$$

we can approximate (4.9) as

$$\mathbf{x}(t+T) = \mathbf{x}(t) + \mathbf{A}\mathbf{x}(t)T + \mathbf{B}\mathbf{u}(t)T \tag{4.11}$$

If we compute $\mathbf{x}(t)$ and $\mathbf{y}(t)$ only at $t = kT$ for $k = 0,\ 1,\ \ldots$, then (4.11) and (4.10) become

$$\mathbf{x}((k+1)T) = (\mathbf{I} + T\mathbf{A})\mathbf{x}(kT) + T\mathbf{B}\mathbf{u}(kT)$$
$$\mathbf{y}(kT) = \mathbf{C}\mathbf{x}(kT) + \mathbf{D}\mathbf{u}(kT)$$

This is a discrete-time state-space equation and can easily be computed on a digital computer. This discretization is the easiest to carry out but yields the least accurate results for the same T. We discuss next a different discretization.

If an input $\mathbf{u}(t)$ is generated by a digital computer followed by a digital-to-analog converter, then $\mathbf{u}(t)$ will be piecewise constant. This situation often arises in computer control of control systems. Let

$$\mathbf{u}(t) = \mathbf{u}(kT) =: \mathbf{u}[k] \qquad \text{for } kT \le t < (k+1)T \tag{4.12}$$

for $k = 0,\ 1,\ 2,\ \ldots$. This input changes values only at discrete-time instants. For this input, the solution of (4.9) still equals (4.5). Computing (4.5) at $t = kT$ and $t = (k+1)T$ yields

$$\mathbf{x}[k] := \mathbf{x}(kT) = e^{\mathbf{A}kT}\mathbf{x}(0) + \int_0^{kT} e^{\mathbf{A}(kT-\tau)}\mathbf{B}\mathbf{u}(\tau)\,d\tau \tag{4.13}$$

and

$$\mathbf{x}[k+1] := \mathbf{x}((k+1)T) = e^{\mathbf{A}(k+1)T}\mathbf{x}(0) + \int_0^{(k+1)T} e^{\mathbf{A}((k+1)T-\tau)}\mathbf{B}\mathbf{u}(\tau)\,d\tau \tag{4.14}$$

Equation (4.14) can be written as

$$\mathbf{x}[k+1] = e^{\mathbf{A}T}\left[e^{\mathbf{A}kT}\mathbf{x}(0) + \int_0^{kT} e^{\mathbf{A}(kT-\tau)}\mathbf{B}\mathbf{u}(\tau)d\tau\right] + \int_{kT}^{(k+1)T} e^{\mathbf{A}(kT+T-\tau)}\mathbf{B}\mathbf{u}(\tau)\,d\tau$$

which becomes, after substituting (4.12) and (4.13) and introducing the new variable $\alpha := kT + T - \tau$,

$$\mathbf{x}[k+1] = e^{\mathbf{A}T}\mathbf{x}[k] + \left(\int_0^T e^{\mathbf{A}\alpha}\,d\alpha\right)\mathbf{B}\mathbf{u}[k]$$

Thus, if an input changes value only at discrete-time instants kT and if we compute only the responses at $t = kT$, then (4.9) and (4.10) become

$$\mathbf{x}[k+1] = \mathbf{A}_d\mathbf{x}[k] + \mathbf{B}_d\mathbf{u}[k] \tag{4.15}$$

$$\mathbf{y}[k] = \mathbf{C}_d\mathbf{x}[k] + \mathbf{D}_d\mathbf{u}[k] \tag{4.16}$$

with

$$\mathbf{A}_d = e^{\mathbf{A}T} \qquad \mathbf{B}_d = \left(\int_0^T e^{\mathbf{A}\tau} d\tau\right)\mathbf{B} \qquad \mathbf{C}_d = \mathbf{C} \qquad \mathbf{D}_d = \mathbf{D} \tag{4.17}$$

This is a discrete-time state-space equation. Note that there is no approximation involved in this derivation and (4.15) yields the exact solution of (4.9) at $t = kT$ if the input is piecewise constant.

We discuss the computation of $\mathbf{B}_d$. Using (3.51), we have

$$\int_0^T \left(\mathbf{I} + \mathbf{A}\tau + \mathbf{A}^2\frac{\tau^2}{2!} + \cdots\right) d\tau$$
$$= T\mathbf{I} + \frac{T^2}{2!}\mathbf{A} + \frac{T^3}{3!}\mathbf{A}^2 + \frac{T^4}{4!}\mathbf{A}^3 + \cdots$$

This power series can be computed recursively as in computing (3.51). If $\mathbf{A}$ is nonsingular, then the series can be written as, using (3.51),

$$\mathbf{A}^{-1}\left(T\mathbf{A} + \frac{T^2}{2!}\mathbf{A}^2 + \frac{T^3}{3!}\mathbf{A}^3 + \cdots + \mathbf{I} - \mathbf{I}\right) = \mathbf{A}^{-1}(e^{\mathbf{A}T} - \mathbf{I})$$

Thus we have

$$\mathbf{B}_d = \mathbf{A}^{-1}(\mathbf{A}_d - \mathbf{I})\mathbf{B} \qquad \text{(if } \mathbf{A} \text{ is nonsingular)} \tag{4.18}$$

Using this formula, we can avoid computing an infinite series.

The MATLAB function `[ad,bd]=c2d(a,b,T)` transforms the continuous-time state equation in (4.9) into the discrete-time state equation in (4.15).

4.2.2 Solution of Discrete-Time Equations

Consider the discrete-time state-space equation

$$\begin{aligned} \mathbf{x}[k+1] &= \mathbf{A}\mathbf{x}[k] + \mathbf{B}\mathbf{u}[k] \\ \mathbf{y}[k] &= \mathbf{C}\mathbf{x}[k] + \mathbf{D}\mathbf{u}[k] \end{aligned} \tag{4.19}$$

where the subscript d has been dropped. It is understood that if the equation is obtained from a continuous-time equation, then the four matrices must be computed from (4.17). The two equations in (4.19) are algebraic equations. Once $\mathbf{x}[0]$ and $\mathbf{u}[k]$, $k = 0, 1, \ldots$, are given, the response can be computed recursively from the equations.

The MATLAB function `dstep` computes unit-step responses of discrete-time state-space equations. It also computes unit-step responses of discrete transfer functions; internally, it first transforms the transfer function into a discrete-time state-space equation by calling `tf2ss`, which will be discussed later, and then uses `dstep`. The function `dlsim`, an acronym for discrete linear simulation, computes responses excited by any input. The function `step` computes unit-step responses of continuous-time state-space equations. Internally, it first uses the function `c2d` to transform a continuous-time state equation into a discrete-time equation and then carries out the computation. If the function `step` is applied to a continuous-time transfer function, then it first uses `tf2ss` to transform the transfer function into a continuous-time state equation and then discretizes it by using `c2d` and then uses `dstep` to compute the response. Similar remarks apply to `lsim`, which computes responses of continuous-time state equations or transfer functions excited by any input.

In order to discuss the general behavior of discrete-time state equations, we will develop a general form of solutions. We compute

$$\mathbf{x}[1] = \mathbf{A}\mathbf{x}[0] + \mathbf{B}\mathbf{u}[0]$$

$$\mathbf{x}[2] = \mathbf{A}\mathbf{x}[1] + \mathbf{B}\mathbf{u}[1] = \mathbf{A}^2\mathbf{x}[0] + \mathbf{A}\mathbf{B}\mathbf{u}[0] + \mathbf{B}\mathbf{u}[1]$$

Proceeding forward, we can readily obtain, for $k > 0$,

$$\mathbf{x}[k] = \mathbf{A}^k\mathbf{x}[0] + \sum_{m=0}^{k-1} \mathbf{A}^{k-1-m}\mathbf{B}\mathbf{u}[m] \tag{4.20}$$

$$\mathbf{y}[k] = \mathbf{C}\mathbf{A}^k\mathbf{x}[0] + \sum_{m=0}^{k-1} \mathbf{C}\mathbf{A}^{k-1-m}\mathbf{B}\mathbf{u}[m] + \mathbf{D}\mathbf{u}[k] \tag{4.21}$$

They are the discrete counterparts of (4.5) and (4.7). Their derivations are considerably simpler than the continuous-time case.

We discuss a general property of the zero-input response $\mathbf{A}^k\mathbf{x}[0]$. Suppose $\mathbf{A}$ has eigenvalue λ_1 with multiplicity 4 and eigenvalue λ_2 with multiplicity 1 and suppose its Jordan form is as shown in the second matrix in (3.39). In other words, λ_1 has index 3 and λ_2 has index 1. Then we have

$$\mathbf{A}^k = \mathbf{Q}\begin{bmatrix} \lambda_1^k & k\lambda_1^{k-1} & k(k-1)\lambda_1^{k-2}/2 & 0 & 0 \\ 0 & \lambda_1^k & k\lambda_1^{k-1} & 0 & 0 \\ 0 & 0 & \lambda_1^k & 0 & 0 \\ 0 & 0 & 0 & \lambda_1^k & 0 \\ 0 & 0 & 0 & 0 & \lambda_2^k \end{bmatrix}\mathbf{Q}^{-1}$$

which implies that every entry of the zero-input response is a linear combination of $\{\lambda_1^k,\ k\lambda_1^{k-1},\ k^2\lambda_1^{k-2},\ \lambda_2^k\}$. These terms are dictated by the eigenvalues and their indices.

If every eigenvalue, simple or repeated, of $\mathbf{A}$ has magnitude less than 1, then every zero-input response will approach zero as $k \to \infty$. If $\mathbf{A}$ has an eigenvalue, simple or repeated, with magnitude larger than 1, then most zero-input responses will grow unbounded as $k \to \infty$. If $\mathbf{A}$ has some eigenvalues with magnitude 1 and all with index 1 and the remaining eigenvalues all have magnitudes less than 1, then no zero-input response will grow unbounded. However, if the index is 2 or higher, then some zero-state response may become unbounded. For example, if $\mathbf{A}$ has eigenvalue 1 with index 2, then $\mathbf{A}^k$ contains the terms $\{1,\ k\}$. If a zero-input response contains the term k, then it will grow unbounded as $k \to \infty$.

4.3 Equivalent State Equations

The example that follows provides a motivation for studying equivalent state equations.

Example 4.3 Consider the network shown in Fig. 4.1. It consists of one capacitor, one inductor, one resistor, and one voltage source. First we select the inductor current x_1 and capacitor voltage x_2 as state variables as shown. The voltage across the inductor is $\dot{x}_1$ and the current through the capacitor is $\dot{x}_2$. The voltage across the resistor is x_2; thus its current is $x_2/1 = x_2$. Clearly we have $x_1 = x_2 + \dot{x}_2$ and $\dot{x}_1 + x_2 - u = 0$. Thus the network is described by the following state equation:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$
$$y = [0 \quad 1]\mathbf{x} \tag{4.22}$$

If, instead, the loop currents $\bar{x}_1$ and $\bar{x}_2$ are chosen as state variables as shown, then the voltage across the inductor is $\dot{\bar{x}}_1$ and the voltage across the resistor is $(\bar{x}_1 - \bar{x}_2) \cdot 1$. From the left-hand-side loop, we have

$$u = \dot{\bar{x}}_1 + \bar{x}_1 - \bar{x}_2 \quad \text{or} \quad \dot{\bar{x}}_1 = -\bar{x}_1 + \bar{x}_2 + u$$

The voltage across the capacitor is the same as the one across the resistor, which is $\bar{x}_1 - \bar{x}_2$. Thus the current through the capacitor is $\dot{\bar{x}}_1 - \dot{\bar{x}}_2$, which equals $\bar{x}_2$ or

$$\dot{\bar{x}}_2 = \dot{\bar{x}}_1 - \bar{x}_2 = -\bar{x}_1 + u$$

Thus the network is also described by the state equation

$$\begin{bmatrix} \dot{\bar{x}}_1 \\ \dot{\bar{x}}_2 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u \tag{4.23}$$
$$y = [1 \quad -1]\bar{\mathbf{x}}$$

The state equations in (4.22) and (4.23) describe the same network; therefore they must be closely related. In fact, they are equivalent as will be established shortly.

Consider the n-dimensional state equation

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}u(t) \end{aligned} \tag{4.24}$$

where $\mathbf{A}$ is an $n \times n$ constant matrix mapping an n-dimensional real space $\mathcal{R}^n$ into itself. The state $\mathbf{x}$ is a vector in $\mathcal{R}^n$ for all t; thus the real space is also called the state space. The state equation in (4.24) can be considered to be associated with the orthonormal basis in (3.8). Now we study the effect on the equation by choosing a different basis.

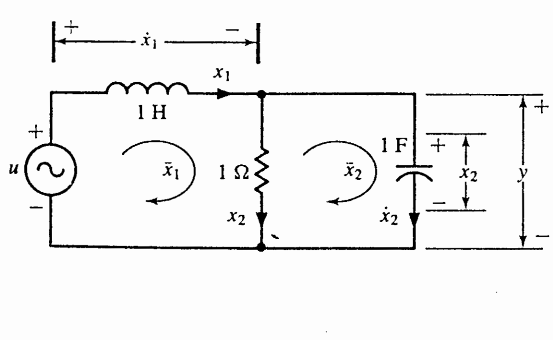

Figure 4.1 Network with two different sets of state variables.

Definition 4.1 *Let* **P** *be an* $n \times n$ *real nonsingular matrix and let* $\bar{\mathbf{x}} = \mathbf{P}\mathbf{x}$. *Then the state equation,*

$$\dot{\bar{\mathbf{x}}}(t) = \bar{\mathbf{A}}\bar{\mathbf{x}}(t) + \bar{\mathbf{B}}\mathbf{u}(t)$$
$$\mathbf{y}(t) = \bar{\mathbf{C}}\bar{\mathbf{x}}(t) + \bar{\mathbf{D}}\mathbf{u}(t) \tag{4.25}$$

where

$$\bar{\mathbf{A}} = \mathbf{P}\mathbf{A}\mathbf{P}^{-1} \qquad \bar{\mathbf{B}} = \mathbf{P}\mathbf{B} \qquad \bar{\mathbf{C}} = \mathbf{C}\mathbf{P}^{-1} \qquad \bar{\mathbf{D}} = \mathbf{D} \tag{4.26}$$

is said to be (algebraically) equivalent *to (4.24) and* $\bar{\mathbf{x}} = \mathbf{P}\mathbf{x}$ *is called an* equivalence transformation.

Equation (4.26) is obtained from (4.24) by substituting $\mathbf{x}(t) = \mathbf{P}^{-1}\bar{\mathbf{x}}(t)$ and $\dot{\mathbf{x}}(t) = \mathbf{P}^{-1}\dot{\bar{\mathbf{x}}}(t)$. In this substitution, we have changed, as in Equation (3.7), the basis vectors of the state space from the orthonormal basis to the columns of $\mathbf{P}^{-1} =: \mathbf{Q}$. Clearly $\mathbf{A}$ and $\bar{\mathbf{A}}$ are similar and $\bar{\mathbf{A}}$ is simply a different representation of $\mathbf{A}$. To be precise, let $\mathbf{Q} = \mathbf{P}^{-1} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \cdots \ \mathbf{q}_n]$. Then the ith column of $\bar{\mathbf{A}}$ is, as discussed in Section 3.4, the representation of $\mathbf{A}\mathbf{q}_i$ with respect to the basis $\{\mathbf{q}_1, \ \mathbf{q}_2, \ \cdots, \ \mathbf{q}_n\}$. From the equation $\bar{\mathbf{B}} = \mathbf{P}\mathbf{B}$ or $\mathbf{B} = \mathbf{P}^{-1}\bar{\mathbf{B}} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \cdots \ \mathbf{q}_n]\bar{\mathbf{B}}$, we see that the ith column of $\bar{\mathbf{B}}$ is the representation of the ith column of $\mathbf{B}$ with respect to the basis $\{\mathbf{q}_1, \ \mathbf{q}_2, \ \cdots, \ \mathbf{q}_n\}$. The matrix $\bar{\mathbf{C}}$ is to be computed from $\mathbf{C}\mathbf{P}^{-1}$. The matrix $\mathbf{D}$, called the *direct transmission part* between the input and output, has nothing to do with the state space and is not affected by the equivalence transformation.

We show that (4.24) and (4.25) have the same set of eigenvalues and the same transfer matrix. Indeed, we have, using $\det(\mathbf{P})\det(\mathbf{P}^{-1}) = 1$,

$$\begin{aligned}\bar{\Delta}(\lambda) &= \det(\lambda\mathbf{I} - \bar{\mathbf{A}}) = \det(\lambda\mathbf{P}\mathbf{P}^{-1} - \mathbf{P}\mathbf{A}\mathbf{P}^{-1}) = \det[\mathbf{P}(\lambda\mathbf{I} - \mathbf{A})\mathbf{P}^{-1}]\\ &= \det(\mathbf{P})\det(\lambda\mathbf{I} - \mathbf{A})\det(\mathbf{P}^{-1}) = \det(\lambda\mathbf{I} - \mathbf{A}) = \Delta(\lambda)\end{aligned}$$

and

$$\begin{aligned}\hat{\bar{\mathbf{G}}}(s) &= \bar{\mathbf{C}}(s\mathbf{I} - \bar{\mathbf{A}})^{-1}\bar{\mathbf{B}} + \bar{\mathbf{D}} = \mathbf{C}\mathbf{P}^{-1}[\mathbf{P}(s\mathbf{I} - \mathbf{A})\mathbf{P}^{-1}]^{-1}\mathbf{P}\mathbf{B} + \mathbf{D}\\ &= \mathbf{C}\mathbf{P}^{-1}\mathbf{P}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{P}^{-1}\mathbf{P}\mathbf{B} + \mathbf{D} = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D} = \hat{\mathbf{G}}(s)\end{aligned}$$

Thus equivalent state equations have the same characteristic polynomial and, consequently, the same set of eigenvalues and same transfer matrix. In fact, all properties of (4.24) are preserved or invariant under any equivalence transformation.

Consider again the network shown in Fig. 4.1, which can be described by (4.22) and (4.23). We show that the two equations are equivalent. From Fig. 4.1, we have $x_1 = \bar{x}_1$. Because the voltage across the resistor is x_2, its current is $x_2/1$ and equals $\bar{x}_1 - \bar{x}_2$. Thus we have

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}$$

or

$$\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{4.27}$$

Note that, for this **P**, its inverse happens to equal itself. It is straightforward to verify that (4.22) and (4.23) are related by the equivalence transformation in (4.26).

The MATLAB function `[ab,bb,cb,db]=ss2ss(a,b,c,d,p)` carries out equivalence transformations.

Two state equations are said to be *zero-state equivalent* if they have the same transfer matrix or

$$\mathbf{D} + \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} = \bar{\mathbf{D}} + \bar{\mathbf{C}}(s\mathbf{I} - \bar{\mathbf{A}})^{-1}\bar{\mathbf{B}}$$

This becomes, after substituting (3.57),

$$\mathbf{D} + \mathbf{CB}s^{-1} + \mathbf{CAB}s^{-2} + \mathbf{CA}^2\mathbf{B}s^{-3} + \cdots = \bar{\mathbf{D}} + \bar{\mathbf{C}}\bar{\mathbf{B}}s^{-1} + \bar{\mathbf{C}}\bar{\mathbf{A}}\bar{\mathbf{B}}s^{-2} + \bar{\mathbf{C}}\bar{\mathbf{A}}^2\bar{\mathbf{B}}s^{-3} + \cdots$$

Thus we have the theorem that follows.

Theorem 4.1

Two linear time-invariant state equations $\{\mathbf{A},\ \mathbf{B},\ \mathbf{C},\ \mathbf{D}\}$ and $\{\bar{\mathbf{A}},\ \bar{\mathbf{B}},\ \bar{\mathbf{C}},\ \bar{\mathbf{D}}\}$ are zero-state equivalent or have the same transfer matrix if and only if $\mathbf{D} = \bar{\mathbf{D}}$ and

$$\mathbf{CA}^m\mathbf{B} = \bar{\mathbf{C}}\bar{\mathbf{A}}^m\bar{\mathbf{B}} \qquad m = 0,\ 1,\ 2,\ \ldots$$

It is clear that (algebraic) equivalence implies zero-state equivalence. In order for two state equations to be equivalent, they must have the same dimension. This is, however, not the case for zero-state equivalence, as the next example shows.

EXAMPLE 4.4 Consider the two networks shown in Fig. 4.2. The capacitor is assumed to have capacitance −1 F. Such a negative capacitance can be realized using an op-amp circuit. For the circuit in Fig. 4.2(a), we have $y(t) = 0.5 \cdot u(t)$ or $\hat{y}(s) = 0.5\hat{u}(s)$. Thus its transfer function is 0.5. To compute the transfer function of the network in Fig. 4.2(b), we may assume the initial voltage across the capacitor to be zero. Because of the symmetry of the four resistors, half of the current will go through each resistor or $i(t) = 0.5u(t)$, where $i(t)$ denotes the right upper resistor's current. Consequently, $y(t) = i(t) \cdot 1 = 0.5u(t)$ and the transfer function also equals 0.5. Thus the two networks, or more precisely their state equations, are zero-state equivalent. This fact can also be verified by using Theorem 4.1. The network in Fig. 4.2(a) is described by the zero-dimensional state equation $y(t) = 0.5u(t)$ or $\mathbf{A} = \mathbf{B} = \mathbf{C} = 0$ and $\mathbf{D} = 0.5$. To

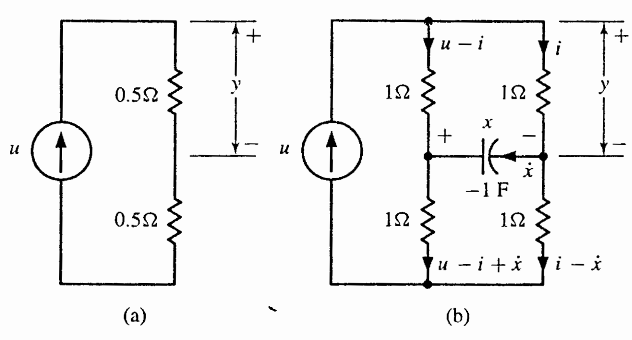

Figure 4.2 Two zero-state equivalent networks.

develop a state equation for the network in Fig. 4.2(b), we assign the capacitor voltage as state variable x with polarity shown. Its current is $\dot{x}$ flowing from the negative to positive polarity because of the negative capacitance. If we assign the right upper resistor's current as $i(t)$, then the right lower resistor's current is $i - \dot{x}$, the left upper resistor's current is $u - i$, and the left lower resistor's current is $u - i + \dot{x}$. The total voltage around the upper right-hand loop is 0:

$$i - x - (u - i) = 0 \quad \text{or} \quad i = 0.5(x + u)$$

which implies

$$y = 1 \cdot i = i = 0.5(x + u)$$

The total voltage around the lower right-hand loop is 0:

$$x + (i - \dot{x}) - (u - i + \dot{x}) = 0$$

which implies

$$2\dot{x} = 2i + x - u = x + u + x - u = 2x$$

Thus the network in Fig. 4.2(b) is described by the one-dimensional state equation

$$\dot{x}(t) = x(t)$$
$$y(t) = 0.5x(t) + 0.5u(t)$$

with $\bar{\mathbf{A}} = 1$, $\bar{\mathbf{B}} = 0$, $\bar{\mathbf{C}} = 0.5$, and $\bar{\mathbf{D}} = 0.5$. We see that $\mathbf{D} = \bar{\mathbf{D}} = 0.5$ and $\mathbf{C}\mathbf{A}^m\mathbf{B} = \bar{\mathbf{C}}\bar{\mathbf{A}}^m\bar{\mathbf{B}} = 0$ for $m = 0, 1, \ldots$. Thus the two equations are zero-state equivalent.

4.3.1 Canonical Forms

MATLAB contains the function `[ab,bb,cb,db,P]=canon(a,b,c,d,'type')`. If `type=companion`, the function will generate an equivalent state equation with $\bar{\mathbf{A}}$ in the companion form in (3.24). This function works only if $\mathbf{Q} := [\mathbf{b}_1 \ \mathbf{A}\mathbf{b}_1 \ \cdots \ \mathbf{A}^{n-1}\mathbf{b}_1]$ is nonsingular, where $\mathbf{b}_1$ is the first column of $\mathbf{B}$. This condition is the same as $\{\mathbf{A}, \mathbf{b}_1\}$ controllable, as we will discuss in Chapter 6. The $\mathbf{P}$ that the function `canon` generates equals $\mathbf{Q}^{-1}$. See the discussion in Section 3.4.

We discuss a different canonical form. Suppose $\mathbf{A}$ has two real eigenvalues and two complex eigenvalues. Because $\mathbf{A}$ has only real coefficients, the two complex eigenvalues must be complex conjugate. Let λ_1, λ_2, $\alpha + j\beta$, and $\alpha - j\beta$ be the eigenvalues and $\mathbf{q}_1$, $\mathbf{q}_2$, $\mathbf{q}_3$, and $\mathbf{q}_4$ be the corresponding eigenvectors, where λ_1, λ_2, α, β, $\mathbf{q}_1$, and $\mathbf{q}_2$ are all real and $\mathbf{q}_4$ equals the complex conjugate of $\mathbf{q}_3$. Define $\mathbf{Q} = [\mathbf{q}_1 \ \mathbf{q}_2 \ \mathbf{q}_3 \ \mathbf{q}_4]$. Then we have

$$\mathbf{J} := \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \alpha + j\beta & 0 \\ 0 & 0 & 0 & \alpha - j\beta \end{bmatrix} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q}$$

Note that $\mathbf{Q}$ and $\mathbf{J}$ can be obtained from `[q,j]=eig(a)` in MATLAB as shown in Examples 3.5 and 3.6. This form is useless in practice but can be transformed into a real matrix by the following similarity transformation

$$\bar{\mathbf{Q}}^{-1}\mathbf{J}\bar{\mathbf{Q}} := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & j & -j \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \alpha + j\beta & 0 \\ 0 & 0 & 0 & \alpha - j\beta \end{bmatrix}$$

$$\cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0.5 & -0.5j \\ 0 & 0 & 0.5 & 0.5j \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \alpha & \beta \\ 0 & 0 & -\beta & \alpha \end{bmatrix} =: \bar{\mathbf{A}}$$

We see that this transformation transforms the complex eigenvalues on the diagonal into a block with the real part of the eigenvalues on the diagonal and the imaginary part on the off-diagonal. This new A-matrix is said to be in *modal* form. The MATLAB function `[ab,bb,cb,db,P]=canon(a,b,c,d, 'modal')` or `canon(a,b,c,d)` with no type specified will yield an equivalent state equation with $\bar{\mathbf{A}}$ in modal form. Note that there is no need to transform **A** into a diagonal form and then to a modal form. The two transformations can be combined into one as

$$\mathbf{P}^{-1} = \mathbf{Q}\bar{\mathbf{Q}} = [\mathbf{q}_1 \ \ \mathbf{q}_2 \ \ \mathbf{q}_3 \ \ \mathbf{q}_4] \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0.5 & -0.5j \\ 0 & 0 & 0.5 & 0.5j \end{bmatrix}$$

$$= [\mathbf{q}_1 \ \mathbf{q}_2 \ \mathrm{Re}(\mathbf{q}_3) \ \mathrm{Im}(\mathbf{q}_3)]$$

where Re and Im stand, respectively, for the real part and imaginary part and we have used in the last equality the fact that $\mathbf{q}_4$ is the complex conjugate of $\mathbf{q}_3$. We give one more example. The modal form of a matrix with real eigenvalue λ_1 and two pairs of distinct complex conjugate eigenvalues $\alpha_i \pm j\beta_i$, for $i = 1, 2$, is

$$\bar{\mathbf{A}} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \alpha_1 & \beta_1 & 0 & 0 \\ 0 & -\beta_1 & \alpha_1 & 0 & 0 \\ 0 & 0 & 0 & \alpha_2 & \beta_2 \\ 0 & 0 & 0 & -\beta_2 & \alpha_2 \end{bmatrix} \tag{4.28}$$

It is block diagonal and can be obtained by the similarity transformation

$$\mathbf{P}^{-1} = [\mathbf{q}_1 \ \ \mathrm{Re}(\mathbf{q}_2) \ \ \mathrm{Im}(\mathbf{q}_2) \ \ \mathrm{Re}(\mathbf{q}_4) \ \ \mathrm{Im}(\mathbf{q}_4)]$$

where $\mathbf{q}_1$, $\mathbf{q}_2$, and $\mathbf{q}_4$ are, respectively, eigenvectors associated with λ_1, $\alpha_1 + j\beta_1$, and $\alpha_2 + j\beta_2$. This form is useful in state-space design.

4.3.2 Magnitude Scaling in Op-Amp Circuits

As discussed in Section 2.3.1, every LTI state equation can be implemented using an op-amp circuit.[1] In actual op-amp circuits, all signals are limited by power supplies. If we use ±15-volt

1. This subsection may be skipped without loss of continuity.

power supplies, then all signals are roughly limited to ± 13 volts. If any signal goes outside the range, the circuit will saturate and will not behave as the state equation dictates. Therefore saturation is an important issue in actual op-amp circuit implementation.

Consider an LTI state equation and suppose all signals must be limited to $\pm M$. For linear systems, if the input magnitude increases by α, so do the magnitudes of all state variables and the output. Thus there must be a limit on input magnitude. Clearly it is desirable to have the admissible input magnitude as large as possible. One way to achieve this is to use an equivalence transformation so that

$$|x_i(t)| \leq |y(t)| \leq M$$

for all i and for all t. The equivalence transformation, however, will not alter the relationship between the input and output; therefore we can use the original state equation to find the input range to achieve $|y(t)| \leq M$. In addition, we can use the same transformation to amplify some state variables to increase visibility or accuracy. This is illustrated in the next example.

Example 4.5 Consider the state equation

$$\dot{\mathbf{x}} = \begin{bmatrix} -0.1 & 2 \\ 0 & -1 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 10 \\ 0.1 \end{bmatrix} u$$

$$y = [0.1 \quad -1]\mathbf{x}$$

Suppose the input is a step function of various magnitude and the equation is to be implemented using an op-amp circuit in which all signals must be limited to ± 10. First we use MATLAB to find its unit-step response. We type

```
a=[-0.1 2;0 -1];b=[10;0.1];c=[0.2 -1];d=0;
[y,x,t]=step(a,b,c,d);
plot(t,y,t,x)
```

which yields the plot in Fig. 4.3(a). We see that $|x_1|_{max} = 100 > |y|_{max} = 20$ and $|x_2| << |y|_{max}$. The state variable x_2 is hardly visible and its largest magnitude is found to be 0.1 by plotting it separately (not shown). From the plot, we see that if $|u(t)| \leq 0.5$, then the output will not saturate but $x_1(t)$ will.

Let us introduce new state variables as

$$\bar{x}_1 = \frac{20}{100} x_1 = 0.2 x_1 \qquad \bar{x}_2 = \frac{20}{0.1} x_2 = 200 x_2$$

With this transformation, the maximum magnitudes of $\bar{x}_1(t)$ and $\bar{x}_2(t)$ will equal $|y|_{max}$. Thus if $y(t)$ does not saturate, neither will all the state variables $\bar{x}_i$. The transformation can be expressed as $\bar{\mathbf{x}} = \mathbf{P}\mathbf{x}$ with

$$\mathbf{P} = \begin{bmatrix} 0.2 & 0 \\ 0 & 200 \end{bmatrix} \qquad \mathbf{P}^{-1} = \begin{bmatrix} 5 & 0 \\ 0 & 0.005 \end{bmatrix}$$

Then its equivalent state equation can readily be computed from (4.26) as

$$\dot{\bar{\mathbf{x}}} = \begin{bmatrix} -0.1 & 0.002 \\ 0 & -1 \end{bmatrix} \bar{\mathbf{x}} + \begin{bmatrix} 2 \\ 20 \end{bmatrix} u$$

$$y = [1 \quad -0.005]\bar{\mathbf{x}}$$

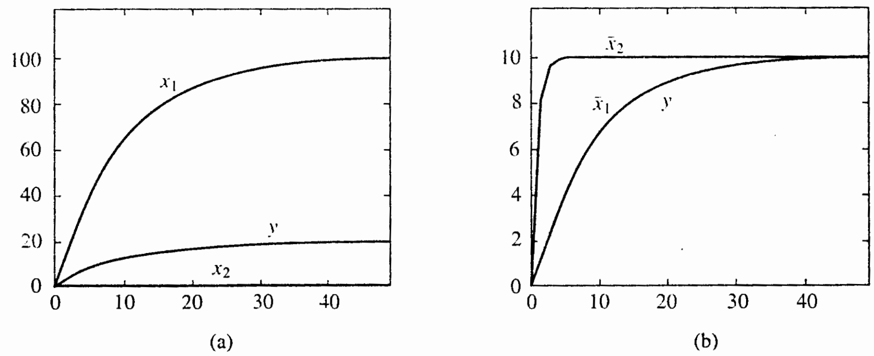

Figure 4.3 Time responses.

Its step responses due to $u(t) = 0.5$ are plotted in Fig. 4.3(b). We see that all signals lie inside the range ± 10 and occupy the full scale. Thus the equivalence state equation is better for op-amp circuit implementation or simulation.

The magnitude scaling is important in using op-amp circuits to implement or simulate continuous-time systems. Although we discuss only step inputs, the idea is applicable to any input. We mention that analog computers are essentially op-amp circuits. Before the advent of digital computers, magnitude scaling in analog computer simulation was carried out by trial and error. With the help of digital computer simulation, the magnitude scaling can now be carried out easily.

4.4 Realizations

Every linear time-invariant (LTI) system can be described by the input–output description

$$\hat{\mathbf{y}}(s) = \hat{\mathbf{G}}(s)\hat{\mathbf{u}}(s)$$

and, if the system is lumped as well, by the state-space equation description

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{aligned} \tag{4.29}$$

If the state equation is known, the transfer matrix can be computed as $\hat{\mathbf{G}}(s) = \mathbf{C}(s\mathbf{I}-\mathbf{A})^{-1}\mathbf{B}+\mathbf{D}$. The computed transfer matrix is unique. Now we study the converse problem, that is, to find a state-space equation from a given transfer matrix. This is called the *realization* problem. This terminology is justified by the fact that, by using the state equation, we can build an op-amp circuit for the transfer matrix.

A transfer matrix $\hat{\mathbf{G}}(s)$ is said to be *realizable* if there exists a finite-dimensional state equation (4.29) or, simply, $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ such that

$$\hat{\mathbf{G}}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}$$

and $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$ is called a *realization* of $\hat{\mathbf{G}}(s)$. An LTI distributed system can be described by a transfer matrix, but not by a finite-dimensional state equation. Thus not every $\hat{\mathbf{G}}(s)$ is realizable. If $\hat{\mathbf{G}}(s)$ is realizable, then it has infinitely many realizations, not necessarily of the same dimension. Thus the realization problem is fairly complex. We study here only the realizability condition. The other issues will be studied in later chapters.

Theorem 4.2

A transfer matrix $\hat{\mathbf{G}}(s)$ is realizable if and only if $\hat{\mathbf{G}}(s)$ is a proper rational matrix.

We use (3.19) to write

$$\hat{\mathbf{G}}_{sp}(s) := \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} = \frac{1}{\det(s\mathbf{I} - \mathbf{A})}\mathbf{C}[\text{Adj}\ (s\mathbf{I} - \mathbf{A})]\mathbf{B} \tag{4.30}$$

If $\mathbf{A}$ is $n \times n$, then det $(s\mathbf{I} - \mathbf{A})$ has degree n. Every entry of Adj $(s\mathbf{I} - \mathbf{A})$ is the determinant of an $(n-1) \times (n-1)$ submatrix of $(s\mathbf{I} - \mathbf{A})$; thus it has at most degree $(n-1)$. Their linear combinations again have at most degree $(n-1)$. Thus we conclude that $\mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}$ is a strictly proper rational matrix. If $\mathbf{D}$ is a nonzero matrix, then $\mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}$ is proper. This shows that if $\hat{\mathbf{G}}(s)$ is realizable, then it is a proper rational matrix. Note that we have

$$\hat{\mathbf{G}}(\infty) = \mathbf{D}$$

Next we show the converse; that is, if $\hat{\mathbf{G}}(s)$ is a $q \times p$ proper rational matrix, then there exists a realization. First we decompose $\hat{\mathbf{G}}(s)$ as

$$\hat{\mathbf{G}}(s) = \hat{\mathbf{G}}(\infty) + \hat{\mathbf{G}}_{sp}(s) \tag{4.31}$$

where $\hat{\mathbf{G}}_{sp}$ is the strictly proper part of $\hat{\mathbf{G}}(s)$. Let

$$d(s) = s^r + \alpha_1 s^{r-1} + \cdots + \alpha_{r-1}s + \alpha_r \tag{4.32}$$

be the least common denominator of all entries of $\hat{\mathbf{G}}_{sp}(s)$. Here we require $d(s)$ to be monic; that is, its leading coefficient is 1. Then $\hat{\mathbf{G}}_{sp}(s)$ can be expressed as

$$\hat{\mathbf{G}}_{sp}(s) = \frac{1}{d(s)}[\mathbf{N}(s)] = \frac{1}{d(s)}\left[\mathbf{N}_1 s^{r-1} + \mathbf{N}_2 s^{r-2} + \cdots + \mathbf{N}_{r-1}s + \mathbf{N}_r\right] \tag{4.33}$$

where $\mathbf{N}_i$ are $q \times p$ constant matrices. Now we claim that the set of equations

$$\dot{\mathbf{x}} = \begin{bmatrix} -\alpha_1\mathbf{I}_p & -\alpha_2\mathbf{I}_p & \cdots & -\alpha_{r-1}\mathbf{I}_p & -\alpha_r\mathbf{I}_p \\ \mathbf{I}_p & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_p & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_p & \mathbf{0} \end{bmatrix}\mathbf{x} + \begin{bmatrix} \mathbf{I}_p \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}\mathbf{u} \tag{4.34}$$

$$\mathbf{y} = \left[\mathbf{N}_1\ \ \mathbf{N}_2\ \ \cdots \mathbf{N}_{r-1}\ \ \mathbf{N}_r\right]\mathbf{x} + \hat{\mathbf{G}}(\infty)\mathbf{u}$$

is a realization of $\hat{\mathbf{G}}(s)$. The matrix $\mathbf{I}_p$ is the $p \times p$ unit matrix and every $\mathbf{0}$ is a $p \times p$ zero matrix. The A-matrix is said to be in block companion form; it consists of r rows and r columns of $p \times p$ matrices; thus the A-matrix has order $rp \times rp$. The B-matrix has order $rp \times p$. Because the C-matrix consists of r number of $\mathbf{N}_i$, each of order $q \times p$, the C-matrix has order $q \times rp$. The realization has dimension rp and is said to be in *controllable canonical form.*

We show that (4.34) is a realization of $\hat{\mathbf{G}}(s)$ in (4.31) and (4.33). Let us define

$$\mathbf{Z} := \begin{bmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \\ \vdots \\ \mathbf{Z}_r \end{bmatrix} := (s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \tag{4.35}$$

where $\mathbf{Z}_i$ is $p \times p$ and $\mathbf{Z}$ is $rp \times p$. Then the transfer matrix of (4.34) equals

$$\mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \hat{\mathbf{G}}(\infty) = \mathbf{N}_1\mathbf{Z}_1 + \mathbf{N}_2\mathbf{Z}_2 + \cdots + \mathbf{N}_r\mathbf{Z}_r + \hat{\mathbf{G}}(\infty) \tag{4.36}$$

We write (4.35) as $(s\mathbf{I} - \mathbf{A})\mathbf{Z} = \mathbf{B}$ or

$$s\mathbf{Z} = \mathbf{A}\mathbf{Z} + \mathbf{B} \tag{4.37}$$

Using the shifting property of the companion form of $\mathbf{A}$, from the second to the last block of equations in (4.37), we can readily obtain

$$s\mathbf{Z}_2 = \mathbf{Z}_1, \quad s\mathbf{Z}_3 = \mathbf{Z}_2, \quad \cdots, \quad s\mathbf{Z}_r = \mathbf{Z}_{r-1}$$

which implies

$$\mathbf{Z}_2 = \frac{1}{s}\mathbf{Z}_1, \quad \mathbf{Z}_3 = \frac{1}{s^2}\mathbf{Z}_1, \quad \cdots, \quad \mathbf{Z}_r = \frac{1}{s^{r-1}}\mathbf{Z}_1$$

Substituting these into the first block of equations in (4.37) yields

$$\begin{aligned} s\mathbf{Z}_1 &= -\alpha_1\mathbf{Z}_1 - \alpha_2\mathbf{Z}_2 - \cdots - \alpha_r\mathbf{Z}_r + \mathbf{I}_p \\ &= -\left(\alpha_1 + \frac{\alpha_2}{s} + \cdots + \frac{\alpha_r}{s^{r-1}}\right)\mathbf{Z}_1 + \mathbf{I}_p \end{aligned}$$

or, using (4.32),

$$\left(s + \alpha_1 + \frac{\alpha_2}{s} + \cdots + \frac{\alpha_r}{s^{r-1}}\right)\mathbf{Z}_1 = \frac{d(s)}{s^{r-1}}\mathbf{Z}_1 = \mathbf{I}_p$$

Thus we have

$$\mathbf{Z}_1 = \frac{s^{r-1}}{d(s)}\mathbf{I}_p, \quad \mathbf{Z}_2 = \frac{s^{r-2}}{d(s)}\mathbf{I}_p, \quad \cdots, \quad \mathbf{Z}_r = \frac{1}{d(s)}\mathbf{I}_p$$

Substituting these into (4.36) yields

$$\mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \hat{\mathbf{G}}(\infty) = \frac{1}{d(s)}[\mathbf{N}_1 s^{r-1} + \mathbf{N}_2 s^{r-2} + \cdots + \mathbf{N}_r] + \hat{\mathbf{G}}(\infty)$$

This equals $\hat{\mathbf{G}}(s)$ in (4.31) and (4.33). This shows that (4.34) is a realization of $\hat{\mathbf{G}}(s)$.

EXAMPLE 4.6 Consider the proper rational matrix

$$\hat{\mathbf{G}}(s) = \begin{bmatrix} \dfrac{4s-10}{2s+1} & \dfrac{3}{s+2} \\ \dfrac{1}{(2s+1)(s+2)} & \dfrac{s+1}{(s+2)^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \dfrac{-12}{2s+1} & \dfrac{3}{s+2} \\ \dfrac{1}{(2s+1)(s+2)} & \dfrac{s+1}{(s+2)^2} \end{bmatrix} \tag{4.38}$$

where we have decomposed $\hat{\mathbf{G}}(s)$ into the sum of a constant matrix and a strictly proper rational matrix $\hat{\mathbf{G}}_{sp}(s)$. The monic least common denominator of $\hat{\mathbf{G}}_{sp}(s)$ is $d(s) = (s+0.5)(s+2)^2 = s^3 + 4.5s^2 + 6s + 2$. Thus we have

$$\hat{\mathbf{G}}_{sp}(s) = \frac{1}{s^3 + 4.5s^2 + 6s + 2} \begin{bmatrix} -6(s+2)^2 & 3(s+2)(s+0.5) \\ 0.5(s+2) & (s+1)(s+0.5) \end{bmatrix}$$

$$= \frac{1}{d(s)} \left(\begin{bmatrix} -6 & 3 \\ 0 & 1 \end{bmatrix} s^2 + \begin{bmatrix} -24 & 7.5 \\ 0.5 & 1.5 \end{bmatrix} s + \begin{bmatrix} -24 & 3 \\ 1 & 0.5 \end{bmatrix} \right)$$

and a realization of (4.38) is

$$\dot{\mathbf{x}} = \begin{bmatrix} -4.5 & 0 & \vdots & -6 & 0 & \vdots & -2 & 0 \\ 0 & -4.5 & \vdots & 0 & -6 & \vdots & 0 & -2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 0 & \vdots & 0 & 0 & \vdots & 0 & 0 \\ 0 & 1 & \vdots & 0 & 0 & \vdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \vdots & 1 & 0 & \vdots & 0 & 0 \\ 0 & 0 & \vdots & 0 & 1 & \vdots & 0 & 0 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \cdots & \cdots \\ 0 & 0 \\ 0 & 0 \\ \cdots & \cdots \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\mathbf{y} = \begin{bmatrix} -6 & 3 & \vdots & -24 & 7.5 & \vdots & -24 & 3 \\ 0 & 1 & \vdots & 0.5 & 1.5 & \vdots & 1 & 0.5 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{4.39}$$

This is a six-dimensional realization.

We discuss a special case of (4.31) and (4.34) in which $p = 1$. To save space, we assume $r = 4$ and $q = 2$. However, the discussion applies to any positive integers r and q. Consider the 2×1 proper rational matrix

$$\hat{\mathbf{G}}(s) = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} + \frac{1}{s^4 + \alpha_1 s^3 + \alpha_2 s^2 + \alpha_3 s + \alpha_4} \cdot \begin{bmatrix} \beta_{11}s^3 + \beta_{12}s^2 + \beta_{13}s + \beta_{14} \\ \beta_{21}s^3 + \beta_{22}s^2 + \beta_{23}s + \beta_{24} \end{bmatrix} \tag{4.40}$$

Then its realization can be obtained directly from (4.34) as

$$\dot{\mathbf{x}} = \begin{bmatrix} -\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u$$
$$\mathbf{y} = \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} \\ \beta_{21} & \beta_{22} & \beta_{23} & \beta_{24} \end{bmatrix} \mathbf{x} + \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} u \tag{4.41}$$

This controllable-canonical-form realization can be read out from the coefficients of $\hat{\mathbf{G}}(s)$ in (4.40).

There are many ways to realize a proper transfer matrix. For example, Problem 4.9 gives a different realization of (4.33) with dimension rq. Let $\hat{\mathbf{G}}_{ci}(s)$ be the ith column of $\hat{\mathbf{G}}(s)$ and let u_i be the ith component of the input vector $\mathbf{u}$. Then $\hat{\mathbf{y}}(s) = \hat{\mathbf{G}}(s)\hat{\mathbf{u}}(s)$ can be expressed as

$$\hat{\mathbf{y}}(s) = \hat{\mathbf{G}}_{c1}(s)\hat{u}_1(s) + \hat{\mathbf{G}}_{c2}(s)\hat{u}_2(s) + \cdots =: \hat{\mathbf{y}}_{c1}(s) + \hat{\mathbf{y}}_{c2}(s) + \cdots$$

as shown in Fig. 4.4(a). Thus we can realize each column of $\hat{\mathbf{G}}(s)$ and then combine them to yield a realization of $\hat{\mathbf{G}}(s)$. Let $\hat{\mathbf{G}}_{ri}(s)$ be the ith row of $\hat{\mathbf{G}}(s)$ and let y_i be the ith component of the output vector $\mathbf{y}$. Then $\hat{\mathbf{y}}(s) = \hat{\mathbf{G}}(s)\hat{\mathbf{u}}(s)$ can be expressed as

$$\hat{y}_i(s) = \hat{\mathbf{G}}_{ri}(s)\hat{\mathbf{u}}(s)$$

as shown in Fig. 4.4(b). Thus we can realize each row of $\hat{\mathbf{G}}(s)$ and then combine them to obtain a realization of $\hat{\mathbf{G}}(s)$. Clearly we can also realize each entry of $\hat{\mathbf{G}}(s)$ and then combine them to obtain a realization of $\hat{\mathbf{G}}(s)$. See Reference [6, pp. 158–160].

The MATLAB function `[a,b,c,d]=tf2ss(num,den)` generates the controllable-canonical-form realization shown in (4.41) for any single-input multiple-output transfer matrix $\hat{\mathbf{G}}(s)$. In its employment, there is no need to decompose $\hat{\mathbf{G}}(s)$ as in (4.31). But we must compute its least common denominator, not necessarily monic. The next example will apply `tf2ss` to each column of $\hat{\mathbf{G}}(s)$ in (4.38) and then combine them to form a realization of $\hat{\mathbf{G}}(s)$.

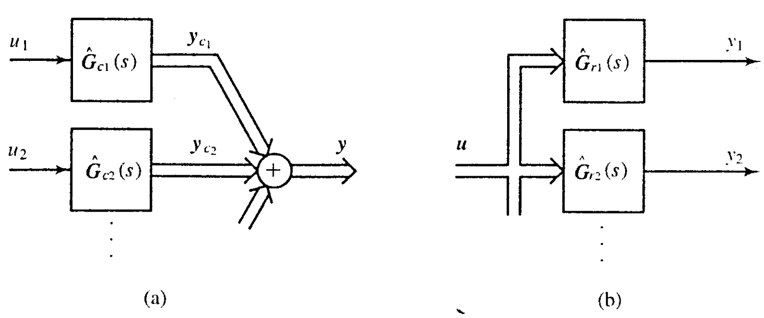

Figure 4.4 Realizations of $\hat{\mathbf{G}}(s)$ by columns and by rows.

EXAMPLE 4.7 Consider the proper rational matrix in (4.38). Its first column is

$$\hat{\mathbf{G}}_{c1}(s) = \begin{bmatrix} \dfrac{4s-10}{2s+1} \\ \dfrac{1}{(2s+1)(s+2)} \end{bmatrix} = \begin{bmatrix} \dfrac{(4s-10)(s+2)}{(2s+1)(s+2)} \\ \dfrac{1}{2s^2+5s+2} \end{bmatrix} = \begin{bmatrix} \dfrac{4s^2-2s-20}{2s^2+5s+2} \\ \dfrac{1}{2s^2+5s+2} \end{bmatrix}$$

Typing

```
n1=[4 -2 -20;0 0 1];d1=[2 5 2]; [a,b,c,d]=tf2ss(n1,d1)
```

yields the following realization for the first column of $\hat{\mathbf{G}}(s)$:

$$\dot{\mathbf{x}}_1 = \mathbf{A}_1\mathbf{x}_1 + \mathbf{b}_1 u_1 = \begin{bmatrix} -2.5 & -1 \\ 1 & 0 \end{bmatrix}\mathbf{x}_1 + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_1$$
$$\mathbf{y}_{c1} = \mathbf{C}_1\mathbf{x}_1 + \mathbf{d}_1 u_1 = \begin{bmatrix} -6 & -12 \\ 0 & 0.5 \end{bmatrix}\mathbf{x}_1 + \begin{bmatrix} 2 \\ 0 \end{bmatrix} u_1 \tag{4.42}$$

Similarly, the function `tf2ss` can generate the following realization for the second column of $\hat{\mathbf{G}}(s)$:

$$\dot{\mathbf{x}}_2 = \mathbf{A}_2\mathbf{x}_2 + \mathbf{b}_2 u_2 = \begin{bmatrix} -4 & -4 \\ 1 & 0 \end{bmatrix}\mathbf{x}_2 + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_2$$
$$\mathbf{y}_{c2} = \mathbf{C}_2\mathbf{x}_2 + \mathbf{d}_2 u_2 = \begin{bmatrix} 3 & 6 \\ 1 & 1 \end{bmatrix}\mathbf{x}_2 + \begin{bmatrix} 0 \\ 0 \end{bmatrix} u_2 \tag{4.43}$$

These two realizations can be combined as

$$\begin{bmatrix} \dot{\mathbf{x}}_1 \\ \dot{\mathbf{x}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 \end{bmatrix}\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{b}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{b}_2 \end{bmatrix}\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$
$$\mathbf{y} = \mathbf{y}_{c1} + \mathbf{y}_{c2} = [\mathbf{C}_1 \ \ \mathbf{C}_2]\mathbf{x} + [\mathbf{d}_1 \ \ \mathbf{d}_2]\mathbf{u}$$

or

$$\dot{\mathbf{x}} = \begin{bmatrix} -2.5 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & -4 & -4 \\ 0 & 0 & 1 & 0 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}\mathbf{u}$$
$$\mathbf{y} = \begin{bmatrix} -6 & -12 & 3 & 6 \\ 0 & 0.5 & 1 & 1 \end{bmatrix}\mathbf{x} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}\mathbf{u} \tag{4.44}$$

This is a different realization of the $\hat{\mathbf{G}}(s)$ in (4.38). This realization has dimension 4, two less than the one in (4.39).

The two state equations in (4.39) and (4.44) are zero-state equivalent because they have the same transfer matrix. They are, however, not algebraically equivalent. More will be said in Chapter 7 regarding realizations. We mention that all discussion in this section, including `tf2ss`, applies without any modification to the discrete-time case

4.5 Solution of Linear Time-Varying (LTV) Equations

Consider the linear time-varying (LTV) state equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) \tag{4.45}$$

$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{u}(t) \tag{4.46}$$

It is assumed that, for every initial state $\mathbf{x}(t_0)$ and any input $\mathbf{u}(t)$, the state equation has a unique solution. A sufficient condition for such an assumption is that every entry of $\mathbf{A}(t)$ is a continuous function of t. Before considering the general case, we first discuss the solutions of $\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t)$ and the reasons why the approach taken in the time-invariant case cannot be used here.

The solution of the time-invariant equation $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ can be extended from the scalar equation $\dot{x} = ax$. The solution of $\dot{x} = ax$ is $x(t) = e^{at}x(0)$ with $d(e^{at})/dt = ae^{at} = e^{at}a$. Similarly, the solution of $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$ is $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{x}(0)$ with

$$\frac{d}{dt}e^{\mathbf{A}t} = \mathbf{A}e^{\mathbf{A}t} = e^{\mathbf{A}t}\mathbf{A}$$

where the commutative property is crucial. Note that, in general, we have $\mathbf{AB} \neq \mathbf{BA}$ and $e^{(\mathbf{A}+\mathbf{B})t} \neq e^{\mathbf{A}t}e^{\mathbf{B}t}$.

The solution of the scalar time-varying equation $\dot{x} = a(t)x$ due to $x(0)$ is

$$x(t) = e^{\int_0^t a(\tau)d\tau}x(0)$$

with

$$\frac{d}{dt}e^{\int_0^t a(\tau)d\tau} = a(t)e^{\int_0^t a(\tau)d\tau} = e^{\int_0^t a(\tau)d\tau}a(t)$$

Extending this to the matrix case becomes

$$\mathbf{x}(t) = e^{\int_0^t \mathbf{A}(\tau)d\tau}\mathbf{x}(0) \tag{4.47}$$

with, using (3.51),

$$e^{\int_0^t \mathbf{A}(\tau)d\tau} = \mathbf{I} + \int_0^t \mathbf{A}(\tau)d\tau + \frac{1}{2}\left(\int_0^t \mathbf{A}(\tau)d\tau\right)\left(\int_0^t \mathbf{A}(s)ds\right) + \cdots$$

This extension, however, is not valid because

$$\begin{aligned}\frac{d}{dt}e^{\int_0^t \mathbf{A}(\tau)d\tau} &= \mathbf{A}(t) + \frac{1}{2}\mathbf{A}(t)\left(\int_0^t \mathbf{A}(s)ds\right) + \frac{1}{2}\left(\int_0^t \mathbf{A}(\tau)d\tau\right)\mathbf{A}(t) + \cdots \\ &\neq \mathbf{A}(t)e^{\int_0^t \mathbf{A}(\tau)d\tau}\end{aligned} \tag{4.48}$$

Thus, in general, (4.47) is not a solution of $\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}$. In conclusion, we cannot extend the solution of scalar time-varying equations to the matrix case and must use a different approach to develop the solution.

Consider

$$\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x} \tag{4.49}$$

where $\mathbf{A}$ is $n \times n$ with continuous functions of t as its entries. Then for every initial state $\mathbf{x}_i(t_0)$, there exists a unique solution $\mathbf{x}_i(t)$, for $i = 1, 2, \ldots, n$. We can arrange these n solutions as $\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_n]$, a square matrix of order n. Because every $\mathbf{x}_i$ satisfies (4.49), we have

$$\dot{\mathbf{X}}(t) = \mathbf{A}(t)\mathbf{X}(t) \tag{4.50}$$

If $\mathbf{X}(t_0)$ is nonsingular or the n initial states are linearly independent, then $\mathbf{X}(t)$ is called a *fundamental matrix* of (4.49). Because the initial states can arbitrarily be chosen, as long as they are linearly independent, the fundamental matrix is not unique.

Example 4.8 Consider the homogeneous equation

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 0 & 0 \\ t & 0 \end{bmatrix} \mathbf{x}(t) \tag{4.51}$$

or

$$\dot{x}_1(t) = 0 \qquad \dot{x}_2(t) = t x_1(t)$$

The solution of $\dot{x}_1(t) = 0$ for $t_0 = 0$ is $x_1(t) = x_1(0)$; the solution of $\dot{x}_2(t) = t x_1(t) = t x_1(0)$ is

$$x_2(t) = \int_0^t \tau x_1(0) d\tau + x_2(0) = 0.5t^2 x_1(0) + x_2(0)$$

Thus we have

$$\mathbf{x}(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow \mathbf{x}(t) = \begin{bmatrix} 1 \\ 0.5t^2 \end{bmatrix}$$

and

$$\mathbf{x}(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \Rightarrow \mathbf{x}(t) = \begin{bmatrix} 1 \\ 0.5t^2 + 2 \end{bmatrix}$$

The two initial states are linearly independent; thus

$$\mathbf{X}(t) = \begin{bmatrix} 1 & 1 \\ 0.5t^2 & 0.5t^2 + 2 \end{bmatrix} \tag{4.52}$$

is a fundamental matrix.

A very important property of the fundamental matrix is that $\mathbf{X}(t)$ is nonsingular for all t. For example, $\mathbf{X}(t)$ in (4.52) has determinant $0.5t^2 + 2 - 0.5t^2 = 2$; thus it is nonsingular for all t. We argue intuitively why this is the case. If $\mathbf{X}(t)$ is singular at some t_1, then there exists a nonzero vector $\mathbf{v}$ such that $\mathbf{x}(t_1) := \mathbf{X}(t_1)\mathbf{v} = \mathbf{0}$, which, in turn, implies $\mathbf{x}(t) := \mathbf{X}(t)\mathbf{v} \equiv \mathbf{0}$ for all t, in particular, at $= t_0$. This is a contradiction. Thus $\mathbf{X}(t)$ is nonsingular for all t.

Definition 4.2 Let $\mathbf{X}(t)$ *be any fundamental matrix of* $\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}$. *Then*

$$\mathbf{\Phi}(t, t_0) := \mathbf{X}(t)\mathbf{X}^{-1}(t_0)$$

is called the state transition matrix *of* $\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}$. *The state transition matrix is also the unique solution of*

$$\frac{\partial}{\partial t}\mathbf{\Phi}(t, t_0) = \mathbf{A}(t)\mathbf{\Phi}(t, t_0) \tag{4.53}$$

with the initial condition $\mathbf{\Phi}(t_0, t_0) = \mathbf{I}$.

Because $\mathbf{X}(t)$ is nonsingular for all t, its inverse is well defined. Equation (4.53) follows directly from (4.50). From the definition, we have the following important properties of the state transition matrix:

$$\mathbf{\Phi}(t, t) = \mathbf{I} \tag{4.54}$$

$$\mathbf{\Phi}^{-1}(t, t_0) = [\mathbf{X}(t)\mathbf{X}^{-1}(t_0)]^{-1} = \mathbf{X}(t_0)\mathbf{X}^{-1}(t) = \mathbf{\Phi}(t_0, t) \tag{4.55}$$

$$\mathbf{\Phi}(t, t_0) = \mathbf{\Phi}(t, t_1)\mathbf{\Phi}(t_1, t_0) \tag{4.56}$$

for every t, t_0, and t_1.

Example 4.9 Consider the homogeneous equation in Example 4.8. Its fundamental matrix was computed as

$$\mathbf{X}(t) = \begin{bmatrix} 1 & 1 \\ 0.5t^2 & 0.5t^2 + 2 \end{bmatrix}$$

Its inverse is, using (3.20),

$$\mathbf{X}^{-1}(t) = \begin{bmatrix} 0.25t^2 + 1 & -0.5 \\ -0.25t^2 & 0.5 \end{bmatrix}$$

Thus the state transition matrix is given by

$$\mathbf{\Phi}(t, t_0) = \begin{bmatrix} 1 & 1 \\ 0.5t^2 & 0.5t^2 + 2 \end{bmatrix}\begin{bmatrix} 0.25t_0^2 + 1 & -0.5 \\ -0.25t_0^2 & 0.5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0.5(t^2 - t_0^2) & 1 \end{bmatrix}$$

It is straightforward to verify that this transition matrix satisfies (4.53) and has the three properties listed in (4.54) through (4.56).

Now we claim that the solution of (4.45) excited by the initial state $\mathbf{x}(t_0) = \mathbf{x}_0$ and the input $\mathbf{u}(t)$ is given by

$$\mathbf{x}(t) = \mathbf{\Phi}(t, t_0)\mathbf{x}_0 + \int_{t_0}^{t} \mathbf{\Phi}(t, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)\,d\tau \tag{4.57}$$

$$= \mathbf{\Phi}(t, t_0)\left[\mathbf{x}_0 + \int_{t_0}^{t} \mathbf{\Phi}(t_0, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)\,d\tau\right] \tag{4.58}$$

where $\mathbf{\Phi}(t, \tau)$ is the state transition matrix of $\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}$. Equation (4.58) follows from (4.57) by using $\mathbf{\Phi}(t, \tau) = \mathbf{\Phi}(t, t_0)\mathbf{\Phi}(t_0, \tau)$. We show that (4.57) satisfies the initial condition and the state equaion. At $t = t_0$, we have

$$\mathbf{x}(t_0) = \mathbf{\Phi}(t_0, t_0)\mathbf{x}_0 + \int_{t_0}^{t_0} \mathbf{\Phi}(t, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)\, d\tau = \mathbf{I}\mathbf{x}_0 + \mathbf{0} = \mathbf{x}_0$$

Thus (4.57) satisfies the initial condition. Using (4.53) and (4.6), we have

$$\begin{aligned}
\frac{d}{dt}\mathbf{x}(t) &= \frac{\partial}{\partial t}\mathbf{\Phi}(t, t_0)\mathbf{x}_0 + \frac{\partial}{\partial t}\int_{t_0}^{t} \mathbf{\Phi}(t, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)\, d\tau \\
&= \mathbf{A}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}_0 + \int_{t_0}^{t} \left(\frac{\partial}{\partial t}\mathbf{\Phi}(t, \tau)\mathbf{B}(\tau)\right) d\tau + \mathbf{\Phi}(t, t)\mathbf{B}(t)\mathbf{u}(t) \\
&= \mathbf{A}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}_0 + \int_{t_0}^{t} \mathbf{A}(t)\mathbf{\Phi}(t, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)\, d\tau + \mathbf{B}(t)\mathbf{u}(t) \\
&= \mathbf{A}(t)\left[\mathbf{\Phi}(t, t_0)\mathbf{x}_0 + \int_{t_0}^{t} \mathbf{\Phi}(t, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)\, d\tau\right] + \mathbf{B}(t)\mathbf{u}(t) \\
&= \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)
\end{aligned}$$

Thus (4.57) is the solution. Substituting (4.57) into (4.46) yields

$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{\Phi}(t, t_0)\mathbf{x}_0 + \mathbf{C}(t)\int_{t_0}^{t} \mathbf{\Phi}(t, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)\, d\tau + \mathbf{D}(t)\mathbf{u}(t) \tag{4.59}$$

If the input is identically zero, then Equation (4.57) reduces to

$$\mathbf{x}(t) = \mathbf{\Phi}(t, t_0)\mathbf{x}_0$$

This is the zero-input response. Thus the state transition matrix governs the unforced propagation of the state vector. If the initial state is zero, then (4.59) reduces to

$$\begin{aligned}
\mathbf{y}(t) &= \mathbf{C}(t)\int_{t_0}^{t} \mathbf{\Phi}(t, \tau)\mathbf{B}(\tau)\mathbf{u}(\tau)\, d\tau + \mathbf{D}(t)\mathbf{u}(t) \\
&= \int_{t_0}^{t} [\mathbf{C}(t)\mathbf{\Phi}(t, \tau)\mathbf{B}(\tau) + \mathbf{D}\delta(t - \tau)]\, \mathbf{u}(\tau)\, d\tau
\end{aligned} \tag{4.60}$$

This is the zero-state response. As discussed in (2.5), the zero-state response can be described by

$$\mathbf{y}(t) = \int_{t_0}^{t} \mathbf{G}(t, \tau)\mathbf{u}(\tau)\, d\tau \tag{4.61}$$

where $\mathbf{G}(t, \tau)$ is the impulse response matrix and is the output at time t excited by an impulse input applied at time τ. Comparing (4.60) and (4.61) yields

$$\begin{aligned}
\mathbf{G}(t, \tau) &= \mathbf{C}(t)\mathbf{\Phi}(t, \tau)\mathbf{B}(\tau) + \mathbf{D}(t)\delta(t - \tau) \\
&= \mathbf{C}(t)\mathbf{X}(t)\mathbf{X}^{-1}(\tau)\mathbf{B}(\tau) + \mathbf{D}(t)\delta(t - \tau)
\end{aligned} \tag{4.62}$$

This relates the input–output and state-space descriptions.

The solutions in (4.57) and (4.59) hinge on solving (4.49) or (4.53). If $\mathbf{A}(t)$ is triangular such as

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} a_{11}(t) & 0 \\ a_{21}(t) & a_{22}(t) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

we can solve the scalar equation $\dot{x}_1(t) = a_{11}(t)x_1(t)$ and then substitute it into

$$\dot{x}_2(t) = a_{22}(t)x_2(t) + a_{21}(t)x_1(t)$$

Because $x_1(t)$ has been solved, the preceding scalar equation can be solved for $x_2(t)$. This is what we did in Example 4.8. If $\mathbf{A}(t)$, such as $\mathbf{A}(t)$ diagonal or constant, has the commutative property

$$\mathbf{A}(t)\left(\int_{t_0}^{t} \mathbf{A}(\tau)\,d\tau\right) = \left(\int_{t_0}^{t} \mathbf{A}(\tau)\,d\tau\right)\mathbf{A}(t)$$

for all t_0 and t, then the solution of (4.53) can be shown to be

$$\mathbf{\Phi}(t, t_0) = e^{\int_{t_0}^{t} \mathbf{A}(\tau)\,d\tau} = \sum_{k=0}^{\infty} \frac{1}{k!}\left(\int_{t_0}^{t} \mathbf{A}(\tau)\,d\tau\right)^k \tag{4.63}$$

For $\mathbf{A}(t)$ constant, (4.63) reduces to

$$\mathbf{\Phi}(t, \tau) = e^{\mathbf{A}(t-\tau)} = \mathbf{\Phi}(t - \tau)$$

and $\mathbf{X}(t) = e^{\mathbf{A}t}$. Other than the preceding special cases, computing state transition matrices is generally difficult.

4.5.1 Discrete-Time Case

Consider the discrete-time state equation

$$\mathbf{x}[k+1] = \mathbf{A}[k]\mathbf{x}[k] + \mathbf{B}[k]\mathbf{u}[k] \tag{4.64}$$

$$\mathbf{y}[k] = \mathbf{C}[k]\mathbf{x}[k] + \mathbf{D}[k]\mathbf{u}[k] \tag{4.65}$$

The set consists of algebraic equations and their solutions can be computed recursively once the initial state $\mathbf{x}[k_0]$ and the input $\mathbf{u}[k]$, for $k \geq k_0$, are given. The situation here is much simpler than the continuous-time case.

As in the continuous-time case, we can define the discrete state transition matrix as the solution of

$$\mathbf{\Phi}[k+1, k_0] = \mathbf{A}[k]\mathbf{\Phi}[k, k_0] \qquad \text{with } \mathbf{\Phi}[k_0, k_0] = \mathbf{I}$$

for $k = k_0, k_0 + 1, \ldots$. This is the discrete counterpart of (4.53) and its solution can be obtained directly as

$$\mathbf{\Phi}[k, k_0] = \mathbf{A}[k-1]\mathbf{A}[k-2]\cdots\mathbf{A}[k_0] \tag{4.66}$$

for $k > k_0$ and $\mathbf{\Phi}[k_0, k_0] = \mathbf{I}$. We discuss a significant difference between the continuous- and discrete-time cases. Because the fundamental matrix in the continuous-time case is nonsingular for all t, the state transition matrix $\mathbf{\Phi}(t, t_0)$ is defined for $t \geq t_0$ and $t < t_0$ and can govern the propagation of the state vector in the positive-time and negative-time directions. In the discrete-time case, the A-matrix can be singular; thus the inverse of $\mathbf{\Phi}[k, k_0]$ may not be defined. Thus $\mathbf{\Phi}[k, k_0]$ is defined only for $k \geq k_0$ and governs the propagation of the state vector in only the positive-time direction. Therefore the discrete counterpart of (4.56) or

$$\mathbf{\Phi}[k, k_0] = \mathbf{\Phi}[k, k_1]\mathbf{\Phi}[k_1, k_0]$$

holds only for $k \geq k_1 \geq k_0$.

Using the discrete state transition matrix, we can express the solutions of (4.64) and (4.65) as, for $k > k_0$,

$$\begin{aligned} \mathbf{x}[k] &= \mathbf{\Phi}[k, k_0]\mathbf{x}_0 + \sum_{m=k_0}^{k-1} \mathbf{\Phi}[k, m+1]\mathbf{B}[m]\mathbf{u}[m] \\ \mathbf{y}[k] &= \mathbf{C}[k]\mathbf{\Phi}[k, k_0]\mathbf{x}_0 + \mathbf{C}[k] \sum_{m=k_0}^{k-1} \mathbf{\Phi}[k, m+1]\mathbf{B}[m]\mathbf{u}[m] + \mathbf{D}[k]\mathbf{u}[k] \end{aligned} \tag{4.67}$$

Their derivations are similar to those of (4.20) and (4.21) and will not be repeated.

If the initial state is zero, Equation (4.67) reduces to

$$\mathbf{y}[k] = \mathbf{C}[k] \sum_{m=k_0}^{k-1} \mathbf{\Phi}[k, m+1]\mathbf{B}[m]\mathbf{u}[m] + \mathbf{D}[k]\mathbf{u}[k] \tag{4.68}$$

for $k > k_0$. This describes the zero-state response of (4.65). If we define $\mathbf{\Phi}[k, m] = \mathbf{0}$ for $k < m$, then (4.68) can be written as

$$\mathbf{y}[k] = \sum_{m=k_0}^{k} \left(\mathbf{C}[k]\mathbf{\Phi}[k, m+1]\mathbf{B}[m] + \mathbf{D}[m]\delta[k-m]\right)\mathbf{u}[m]$$

where the impulse sequence $\delta[k-m]$ equals 1 if $k = m$ and 0 if $k \neq m$. Comparing this with the multivariable version of (2.34), we have

$$\mathbf{G}[k, m] = \mathbf{C}[k]\mathbf{\Phi}[k, m+1]\mathbf{B}[m] + \mathbf{D}[m]\delta[k-m]$$

for $k \geq m$. This relates the impulse response sequence and the state equation and is the discrete counterpart of (4.62).

4.6 Equivalent Time-Varying Equations

This section extends the equivalent state equations discussed in Section 4.3 to the time-varying case. Consider the n-dimensional linear time-varying state equation

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}(t)\mathbf{x} + \mathbf{B}(t)\mathbf{u} \\ \mathbf{y} &= \mathbf{C}(t)\mathbf{x} + \mathbf{D}(t)\mathbf{u} \end{aligned} \tag{4.69}$$

Let $\mathbf{P}(t)$ be an $n \times n$ matrix. It is assumed that $\mathbf{P}(t)$ is nonsingular and both $\mathbf{P}(t)$ and $\dot{\mathbf{P}}(t)$ are continuous for all t. Let $\bar{\mathbf{x}} = \mathbf{P}(t)\mathbf{x}$. Then the state equation

$$\begin{aligned} \dot{\bar{\mathbf{x}}} &= \bar{\mathbf{A}}(t)\bar{\mathbf{x}} + \bar{\mathbf{B}}(t)\mathbf{u} \\ \mathbf{y} &= \bar{\mathbf{C}}(t)\bar{\mathbf{x}} + \bar{\mathbf{D}}(t)\mathbf{u} \end{aligned} \tag{4.70}$$

where

$$\begin{aligned} \bar{\mathbf{A}}(t) &= [\mathbf{P}(t)\mathbf{A}(t) + \dot{\mathbf{P}}(t)]\mathbf{P}^{-1}(t) \\ \bar{\mathbf{B}}(t) &= \mathbf{P}(t)\mathbf{B}(t) \\ \bar{\mathbf{C}}(t) &= \mathbf{C}(t)\mathbf{P}^{-1}(t) \\ \bar{\mathbf{D}}(t) &= \mathbf{D}(t) \end{aligned}$$

is said to be (algebraically) equivalent to (4.69) and $\mathbf{P}(t)$ is called an *(algebraic) equivalence* transformation.

Equation (4.70) is obtained from (4.69) by substituting $\bar{\mathbf{x}} = \mathbf{P}(t)\mathbf{x}$ and $\dot{\bar{\mathbf{x}}} = \dot{\mathbf{P}}(t)\mathbf{x}+\mathbf{P}(t)\dot{\mathbf{x}}$. Let $\mathbf{X}$ be a fundamental matrix of (4.69). Then we claim that

$$\bar{\mathbf{X}}(t) := \mathbf{P}(t)\mathbf{X}(t) \tag{4.71}$$

is a fundamental matrix of (4.70). By definition, $\dot{\mathbf{X}}(t) = \mathbf{A}(t)\mathbf{X}(t)$ and $\mathbf{X}(t)$ is nonsingular for all t. Because the rank of a matrix will not change by multiplying a nonsingular matrix, the matrix $\mathbf{P}(t)\mathbf{X}(t)$ is also nonsingular for all t. Now we show that $\mathbf{P}(t)\mathbf{X}(t)$ satisfies the equation $\dot{\bar{\mathbf{x}}} = \bar{\mathbf{A}}(t)\bar{\mathbf{x}}$. Indeed, we have

$$\begin{aligned} \frac{d}{dt}[\mathbf{P}(t)\mathbf{X}(t)] &= \dot{\mathbf{P}}(t)\mathbf{X}(t) + \mathbf{P}(t)\dot{\mathbf{X}}(t) = \dot{\mathbf{P}}(t)\mathbf{X}(t) + \mathbf{P}(t)\mathbf{A}(t)\mathbf{X}(t) \\ &= [\dot{\mathbf{P}}(t) + \mathbf{P}(t)\mathbf{A}(t)][\mathbf{P}^{-1}(t)\mathbf{P}(t)]\mathbf{X}(t) = \bar{\mathbf{A}}(t)[\mathbf{P}(t)\mathbf{X}(t)] \end{aligned}$$

Thus $\mathbf{P}(t)\mathbf{X}(t)$ is a fundamental matrix of $\dot{\bar{\mathbf{x}}}(t) = \bar{\mathbf{A}}(t)\bar{\mathbf{x}}(t)$.

Theorem 4.3

Let $\mathbf{A}_o$ be an arbitrary constant matrix. Then there exists an equivalence transformation that transforms (4.69) into (4.70) with $\bar{\mathbf{A}}(t) = \mathbf{A}_o$.

Proof: Let $\mathbf{X}(t)$ be a fundamental matrix of $\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}$. The differentiation of $\mathbf{X}^{-1}(t)$ $\mathbf{X}(t) = \mathbf{I}$ yields

$$\dot{\mathbf{X}}^{-1}(t)\mathbf{X}(t) + \mathbf{X}^{-1}\dot{\mathbf{X}}(t) = \mathbf{0}$$

which implies

$$\dot{\mathbf{X}}^{-1}(t) = -\mathbf{X}^{-1}(t)\mathbf{A}(t)\mathbf{X}(t)\mathbf{X}^{-1}(t) = -\mathbf{X}^{-1}(t)\mathbf{A}(t) \tag{4.72}$$

Because $\bar{\mathbf{A}}(t) = \mathbf{A}_o$ is a constant matrix, $\bar{\mathbf{X}}(t) = e^{\mathbf{A}_o t}$ is a fundamental matrix of $\dot{\bar{\mathbf{x}}} = \bar{\mathbf{A}}(t)\bar{\mathbf{x}} = \mathbf{A}_o\bar{\mathbf{x}}$. Following (4.71), we define

$$\mathbf{P}(t) := \bar{\mathbf{X}}(t)\mathbf{X}^{-1}(t) = e^{\mathbf{A}_o t}\mathbf{X}^{-1}(t) \tag{4.73}$$

and compute

$$\bar{\mathbf{A}}(t) = [\mathbf{P}(t)\mathbf{A}(t) + \dot{\mathbf{P}}(t)]\mathbf{P}^{-1}(t)$$
$$= [e^{\mathbf{A}_o t}\mathbf{X}^{-1}(t)\mathbf{A}(t) + \mathbf{A}_o e^{\mathbf{A}_o t}\mathbf{X}^{-1}(t) + e^{\mathbf{A}_o t}\dot{\mathbf{X}}^{-1}(t)]\mathbf{X}(t)e^{-\mathbf{A}_o t}$$

which becomes, after substituting (4.72),

$$\bar{\mathbf{A}}(t) = \mathbf{A}_o e^{\mathbf{A}_o t}\mathbf{X}^{-1}(t)\mathbf{X}(t)e^{-\mathbf{A}_o t} = \mathbf{A}_o$$

This establishes the theorem. Q.E.D.

If $\mathbf{A}_o$ is chosen as a zero matrix, then $\mathbf{P}(t) = \mathbf{X}^{-1}(t)$ and (4.70) reduces to

$$\bar{\mathbf{A}}(t) = \mathbf{0} \qquad \bar{\mathbf{B}}(t) = \mathbf{X}^{-1}(t)\mathbf{B}(t) \qquad \bar{\mathbf{C}}(t) = \mathbf{C}(t)\mathbf{X}(t) \qquad \bar{\mathbf{D}}(t) = \mathbf{D}(t) \qquad (4.74)$$

The block diagrams of (4.69) with $\mathbf{A}(t) \neq \mathbf{0}$ and $\mathbf{A}(t) = \mathbf{0}$ are plotted in Fig. 4.5. The block diagram with $\mathbf{A}(t) = \mathbf{0}$ has no feedback and is considerably simpler. Every time-varying state equation can be transformed into such a block diagram. However, in order to do so, we must know its fundamental matrix.

The impulse response matrix of (4.69) is given in (4.62). The impulse response matrix of (4.70) is, using (4.71) and (4.72),

$$\bar{\mathbf{G}}(t, \tau) = \bar{\mathbf{C}}(t)\bar{\mathbf{X}}(t)\bar{\mathbf{X}}^{-1}(\tau)\bar{\mathbf{B}}(\tau) + \bar{\mathbf{D}}(t)\delta(t - \tau)$$

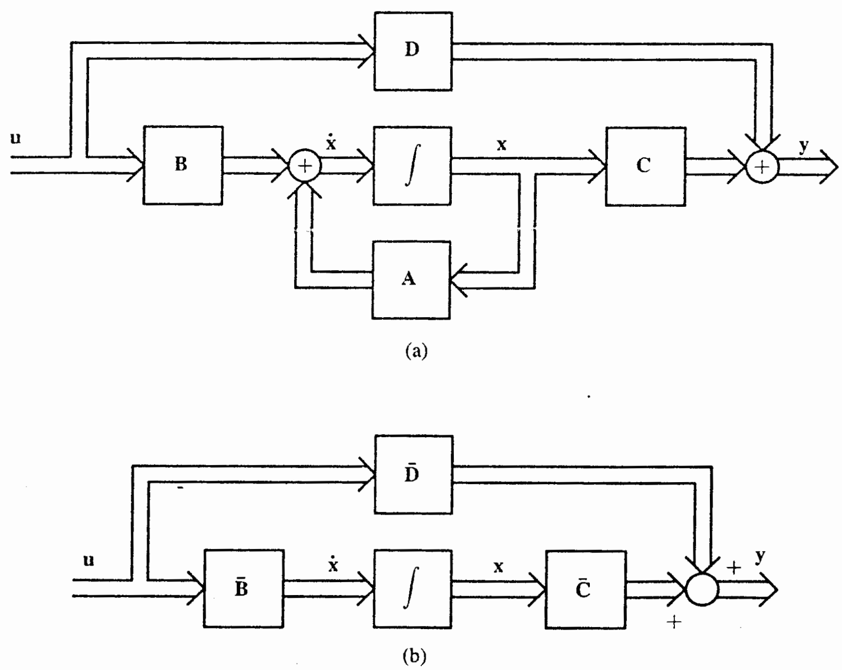

Figure 4.5 Block daigrams with feedback and without feedback.

$$= \mathbf{C}(t)\mathbf{P}^{-1}(t)\mathbf{P}(t)\mathbf{X}(t)\mathbf{X}^{-1}(\tau)\mathbf{P}^{-1}(\tau)\mathbf{P}(\tau)\mathbf{B}(\tau) + \mathbf{D}(t)\delta(t-\tau)$$
$$= \mathbf{C}(t)\mathbf{X}(t)\mathbf{X}^{-1}(\tau)\mathbf{B}(\tau) + \mathbf{D}(t)\delta(t-\tau) = \mathbf{G}(t,\tau)$$

Thus the impulse response matrix is invariant under any equivalence transformation. The property of the A-matrix, however, may not be preserved in equivalence transformations. For example, every A-matrix can be transformed, as shown in Theorem 4.3, into a constant or a zero matrix. Clearly the zero matrix does not have any property of $\mathbf{A}(t)$. In the time-invariant case, equivalence transformations will preserve all properties of the original state equation. Thus the equivalence transformation in the time-invariant case is not a special case of the time-varying case.

> ***Definition 4.3*** *A matrix* $\mathbf{P}(t)$ *is called a* Lyapunov transformation *if* $\mathbf{P}(t)$ *is nonsingular,* $\mathbf{P}(t)$ *and* $\dot{\mathbf{P}}(t)$ *are continuous, and* $\mathbf{P}(t)$ *and* $\mathbf{P}^{-1}(t)$ *are bounded for all* t. *Equations (4.69) and (4.70) are said to be* Lyapunov equivalent *if* $\mathbf{P}(t)$ *is a Lyapunov transformation.*

It is clear that if $\mathbf{P}(t) = \mathbf{P}$ is a constant matrix, then it is a Lyapunov transformation. Thus the (algebraic) transformation in the time-invariant case is a special case of the Lyapunov transformation. If $\mathbf{P}(t)$ is required to be a Lyapunov transformation, then Theorem 4.3 does not hold in general. In other words, not every time-varying state equation can be Lyapunov equivalent to a state equation with a constant A-matrix. However, this is true if $\mathbf{A}(t)$ is periodic.

Periodic state equations Consider the linear time-varying state equation in (4.69). It is assumed that

$$\mathbf{A}(t+T) = \mathbf{A}(t)$$

for all t and for some positive constant T. That is, $\mathbf{A}(t)$ is periodic with period T. Let $\mathbf{X}(t)$ be a fundamental matrix of $\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}$ or $\dot{\mathbf{X}}(t) = \mathbf{A}(t)\mathbf{X}(t)$ with $\mathbf{X}(0)$ nonsingular. Then we have

$$\dot{\mathbf{X}}(t+T) = \mathbf{A}(t+T)\mathbf{X}(t+T) = \mathbf{A}(t)\mathbf{X}(t+T)$$

Thus $\mathbf{X}(t+T)$ is also a fundamental matrix. Furthermore, it can be expressed as

$$\mathbf{X}(t+T) = \mathbf{X}(t)\mathbf{X}^{-1}(0)\mathbf{X}(T) \tag{4.75}$$

This can be verified by direct substitution. Let us define $\mathbf{Q} = \mathbf{X}^{-1}(0)\mathbf{X}(T)$. It is a constant nonsingular matrix. For this $\mathbf{Q}$ there exists a constant matrix $\bar{\mathbf{A}}$ such that $e^{\bar{\mathbf{A}}T} = \mathbf{Q}$ (Problem 3.24). Thus (4.75) can be written as

$$\mathbf{X}(t+T) = \mathbf{X}(t)e^{\bar{\mathbf{A}}T} \tag{4.76}$$

Define

$$\mathbf{P}(t) := e^{\bar{\mathbf{A}}t}\mathbf{X}^{-1}(t) \tag{4.77}$$

We show that $\mathbf{P}(t)$ is periodic with period T:

$$\mathbf{P}(t+T) = e^{\bar{\mathbf{A}}(t+T)}\mathbf{X}^{-1}(t+T) = e^{\bar{\mathbf{A}}t}e^{\bar{\mathbf{A}}T}[e^{-\bar{\mathbf{A}}T}\mathbf{X}^{-1}(t)]$$
$$= e^{\bar{\mathbf{A}}t}\mathbf{X}^{-1}(t) = \mathbf{P}(t)$$

Theorem 4.4

Consider (4.69) with $\mathbf{A}(t) = \mathbf{A}(t+T)$ for all t and some $T > 0$. Let $\mathbf{X}(t)$ be a fundamental matrix of $\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}$. Let $\bar{\mathbf{A}}$ be the constant matrix computed from $e^{\bar{\mathbf{A}}T} = \mathbf{X}^{-1}(0)\mathbf{X}(T)$. Then (4.69) is Lyapunov equivalent to

$$\dot{\bar{\mathbf{x}}}(t) = \bar{\mathbf{A}}\bar{\mathbf{x}}(t) + \mathbf{P}(t)\mathbf{B}(t)\mathbf{u}(t)$$
$$\bar{\mathbf{y}}(t) = \mathbf{C}(t)\mathbf{P}^{-1}(t)\bar{\mathbf{x}}(t) + \mathbf{D}(t)\mathbf{u}(t)$$

where $\mathbf{P}(t) = e^{\bar{\mathbf{A}}t}\mathbf{X}^{-1}(t)$.

The matrix $\mathbf{P}(t)$ in (4.77) satisfies all conditions in Definition 4.3; thus it is a Lyapunov transformation. The rest of the theorem follows directly from Theorem 4.3. The homogeneous part of Theorem 4.4 is the *theory of Floquet*. It states that if $\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}$ and if $\mathbf{A}(t+T) = \mathbf{A}(t)$ for all t, then its fundamental matrix is of the form $\mathbf{P}^{-1}(t)e^{\bar{\mathbf{A}}t}$, where $\mathbf{P}^{-1}(t)$ is a periodic function. Furthermore, $\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}$ is Lyapunov equivalent to $\dot{\bar{\mathbf{x}}} = \bar{\mathbf{A}}\bar{\mathbf{x}}$.

4.7 Time-Varying Realizations

We studied in Section 4.4 the realization problem for linear time-invariant systems. In this section, we study the corresponding problem for linear time-varying systems. The Laplace transform cannot be used here; therefore we study the problem directly in the time domain.

Every linear time-varying system can be described by the input–output description

$$\mathbf{y}(t) = \int_{t_0}^{t} \mathbf{G}(t, \tau)\mathbf{u}(\tau)\, d\tau$$

and, if the system is lumped as well, by the state equation

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{u}(t) \end{aligned} \tag{4.78}$$

If the state equation is available, the impulse response matrix can be computed from

$$\mathbf{G}(t, \tau) = \mathbf{C}(t)\mathbf{X}(t)\mathbf{X}^{-1}(\tau)\mathbf{B}(\tau) + \mathbf{D}(t)\delta(t-\tau) \qquad \text{for } t \geq \tau \tag{4.79}$$

where $\mathbf{X}(t)$ is a fundamental matrix of $\dot{\mathbf{x}} = \mathbf{A}(t)\mathbf{x}$. The converse problem is to find a state equation from a given impulse response matrix. An impulse response matrix $\mathbf{G}(t, \tau)$ is said to be *realizable* if there exists $\{\mathbf{A}(t),\ \mathbf{B}(t),\ \mathbf{C}(t),\ \mathbf{D}(t)\}$ to meet (4.79).

Theorem 4.5

A $q \times p$ impulse response matrix $\mathbf{G}(t, \tau)$ is realizable if and only if $\mathbf{G}(t, \tau)$ can be decomposed as

$$\mathbf{G}(t, \tau) = \mathbf{M}(t)\mathbf{N}(\tau) + \mathbf{D}(t)\delta(t-\tau) \tag{4.80}$$

for all $t \geq \tau$, where $\mathbf{M}$, $\mathbf{N}$, and $\mathbf{D}$ are, respectively, $q \times n$, $n \times p$, and $q \times p$ matrices for some integer n.

Proof: If $\mathbf{G}(t, \tau)$ is realizable, there exists a realization that meets (4.79). Identifying $\mathbf{M}(t) = \mathbf{C}(t)\mathbf{X}(t)$ and $\mathbf{N}(\tau) = \mathbf{X}^{-1}(\tau)\mathbf{B}(\tau)$ establishes the necessary part of the theorem.

If $\mathbf{G}(t, \tau)$ can be decomposed as in (4.80), then the n-dimensional state equation

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{N}(t)\mathbf{u}(t)\\ \mathbf{y}(t) &= \mathbf{M}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{u}(t)\end{aligned} \tag{4.81}$$

is a realization. Indeed, a fundamental matrix of $\dot{\mathbf{x}} = \mathbf{0} \cdot \mathbf{x}$ is $\mathbf{X}(t) = \mathbf{I}$. Thus the impulse response matrix of (4.81) is

$$\mathbf{M}(t)\mathbf{I} \cdot \mathbf{I}^{-1}\mathbf{N}(\tau) + \mathbf{D}(t)\delta(t - \tau)$$

which equals $\mathbf{G}(t, \tau)$. This shows the sufficiency of the theorem. Q.E.D.

Although Theorem 4.5 can also be applied to time-invariant systems, the result is not useful in practical implementation, as the next example illustrates.

EXAMPLE 4.10 Consider $g(t) = te^{\lambda t}$ or

$$g(t, \tau) = g(t - \tau) = (t - \tau)e^{\lambda(t-\tau)}$$

It is straightforward to verify

$$g(t - \tau) = [e^{\lambda t} \quad te^{\lambda t}]\begin{bmatrix} -\tau e^{-\lambda \tau} \\ e^{-\lambda \tau} \end{bmatrix}$$

Thus the two-dimensional time-varying state equation

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\mathbf{x} + \begin{bmatrix} -te^{-\lambda t} \\ e^{-\lambda t} \end{bmatrix}u(t)\\ y(t) &= [e^{\lambda t} \quad te^{\lambda t}]\mathbf{x}(t)\end{aligned} \tag{4.82}$$

is a realization of the impulse response $g(t) = te^{\lambda t}$.

The Laplace transform of the impulse response is

$$\hat{g}(s) = \mathcal{L}[te^{\lambda t}] = \frac{1}{(s - \lambda)^2} = \frac{1}{s^2 - 2\lambda s + \lambda^2}$$

Using (4.41), we can readily obtain

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \begin{bmatrix} 2\lambda & -\lambda^2 \\ 1 & 0 \end{bmatrix}\mathbf{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix}u(t)\\ y(t) &= [0 \quad 1]\mathbf{x}(t)\end{aligned} \tag{4.83}$$

This LTI state equation is a different realization of the same impulse response.This realization is clearly more desirable because it can readily be implemented using an op-amp circuit. The implementation of (4.82) is much more difficult in practice.

PROBLEMS

4.1 An oscillation can be generated by

$$\dot{\mathbf{x}} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{x}$$

Show that its solution is

$$\mathbf{x}(t) = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix} \mathbf{x}(0)$$

4.2 Use two different methods to find the unit-step response of

$$\dot{\mathbf{x}} = \begin{bmatrix} 0 & 1 \\ -2 & -2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u$$

$$y = [2 \;\; 3]\mathbf{x}$$

4.3 Discretize the state equation in Problem 4.2 for $T = 1$ and $T = \pi$.

4.4 Find the companion-form and modal-form equivalent equations of

$$\dot{\mathbf{x}} = \begin{bmatrix} -2 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & -2 & -2 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} u$$

$$y = [1 \;\; -1 \;\; 0]\mathbf{x}$$

4.5 Find an equivalent state equation of the equation in Problem 4.4 so that all state variables have their largest magnitudes roughly equal to the largest magnitude of the output. If all signals are required to lie inside ± 10 volts and if the input is a step function with magnitude a, what is the permissible largest a?

4.6 Consider

$$\dot{\mathbf{x}} = \begin{bmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{bmatrix} \mathbf{x} + \begin{bmatrix} b_1 \\ \bar{b}_1 \end{bmatrix} u \qquad y = [c_1 \;\; \bar{c}_1]\mathbf{x}$$

where the overbar denotes complex conjugate. Verify that the equation can be transformed into

$$\dot{\bar{\mathbf{x}}} = \bar{\mathbf{A}}\bar{\mathbf{x}} + \bar{\mathbf{b}}u \qquad y = \bar{\mathbf{c}}\bar{\mathbf{x}}$$

with

$$\bar{\mathbf{A}} = \begin{bmatrix} 0 & 1 \\ -\lambda\bar{\lambda} & \lambda + \bar{\lambda} \end{bmatrix} \qquad \bar{\mathbf{b}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \bar{\mathbf{c}}_1 = [-2\mathrm{Re}(\bar{\lambda} b_1 c_1) \;\; 2\mathrm{Re}(b_1 c_1)]$$

by using the transformation $\mathbf{x} = \mathbf{Q}\bar{\mathbf{x}}$ with

$$\mathbf{Q}_1 = \begin{bmatrix} -\bar{\lambda} b_1 & b_1 \\ -\lambda \bar{b}_1 & \bar{b}_1 \end{bmatrix}$$

4.7 Verify that the Jordan-form equation

$$\dot{\mathbf{x}} = \begin{bmatrix} \lambda & 1 & 0 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 & 0 \\ 0 & 0 & \lambda & 0 & 0 & 0 \\ 0 & 0 & 0 & \bar{\lambda} & 1 & 0 \\ 0 & 0 & 0 & 0 & \bar{\lambda} & 1 \\ 0 & 0 & 0 & 0 & 0 & \bar{\lambda} \end{bmatrix} \mathbf{x} + \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \bar{b}_1 \\ \bar{b}_2 \\ \bar{b}_3 \end{bmatrix} u$$

$$y = [c_1 \ c_2 \ c_3 \ \bar{c}_1 \ \bar{c}_2 \ \bar{c}_3]\mathbf{x}$$

can be transformed into

$$\dot{\bar{\mathbf{x}}} = \begin{bmatrix} \bar{\mathbf{A}} & \mathbf{I}_2 & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{A}} & \mathbf{I}_2 \\ \mathbf{0} & \mathbf{0} & \bar{\mathbf{A}} \end{bmatrix} \bar{\mathbf{x}} + \begin{bmatrix} \bar{\mathbf{b}} \\ \bar{\mathbf{b}} \\ \bar{\mathbf{b}} \end{bmatrix} u \qquad y = [\bar{\mathbf{c}}_1 \ \bar{\mathbf{c}}_2 \ \bar{\mathbf{c}}_3]\bar{\mathbf{x}}$$

where $\bar{\mathbf{A}}$, $\bar{\mathbf{b}}$, and $\bar{\mathbf{c}}_i$ are defined in Problem 4.6 and $\mathbf{I}_2$ is the unit matrix of order 2. [*Hint*: Change the order of the state variables from $[x_1 \ x_2 \ x_3 \ x_4 \ x_5 \ x_6]'$ to $[x_1 \ x_4 \ x_2 \ x_5 \ x_3 \ x_6]'$ and then apply the equivalence transformation $\mathbf{x} = \mathbf{Q}\bar{\mathbf{x}}$ with $\mathbf{Q} = \text{diag}(\mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3)$.]

4.8 Are the two sets of state equations

$$\dot{\mathbf{x}} = \begin{bmatrix} 2 & 1 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{x} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} u \qquad y = [1 \ -1 \ 0]\mathbf{x}$$

and

$$\dot{\mathbf{x}} = \begin{bmatrix} 2 & 1 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & -1 \end{bmatrix} \mathbf{x} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} u \qquad y = [1 \ -1 \ 0]\mathbf{x}$$

equivalent? Are they zero-state equivalent?

4.9 Verify that the transfer matrix in (4.33) has the following realization:

$$\dot{\mathbf{x}} = \begin{bmatrix} -\alpha_1\mathbf{I}_q & \mathbf{I}_q & \mathbf{0} & \cdots & \mathbf{0} \\ -\alpha_2\mathbf{I}_q & \mathbf{0} & \mathbf{I}_q & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & & \vdots \\ -\alpha_{r-1}\mathbf{I}_q & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{I}_q \\ -\alpha_r\mathbf{I}_q & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{N}_1 \\ \mathbf{N}_2 \\ \vdots \\ \mathbf{N}_{r-1} \\ \mathbf{N}_r \end{bmatrix} \mathbf{u}$$

$$\mathbf{y} = [\mathbf{I}_q \ \mathbf{0} \ \mathbf{0} \ \cdots \ \mathbf{0}]\mathbf{x}$$

This is called the *observable canonical form realization* and has dimension rq. It is dual to (4.34).

4.10 Consider the 1×2 proper rational matrix

$$\hat{\mathbf{G}}(s) = [d_1 \ d_2] + \frac{1}{s^4 + \alpha_1 s^3 + \alpha_2 s^2 + \alpha_3 s + \alpha_4}$$

$$\times\ [\beta_{11}s^3 + \beta_{21}s^2 + \beta_{31}s + \beta_{41} \quad \beta_{12}s^3 + \beta_{22}s^2 + \beta_{32}s + \beta_{42}]$$

Show that its observable canonical form realization can be reduced from Problem 4.9 as

$$\dot{\mathbf{x}} = \begin{bmatrix} -\alpha_1 & 1 & 0 & 0 \\ -\alpha_2 & 0 & 1 & 0 \\ -\alpha_3 & 0 & 0 & 1 \\ -\alpha_4 & 0 & 0 & 0 \end{bmatrix} \mathbf{x} + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \\ \beta_{31} & \beta_{32} \\ \beta_{41} & \beta_{42} \end{bmatrix} \mathbf{u}$$

$$y = [1 \ \ 0 \ \ 0 \ \ 0]\mathbf{x} + [d_1 \ \ d_2]\mathbf{u}$$

4.11 Find a realization for the proper rational matrix

$$\hat{\mathbf{G}}(s) = \begin{bmatrix} \dfrac{2}{s+1} & \dfrac{2s-3}{(s+1)(s+2)} \\ \dfrac{s-2}{s+1} & \dfrac{s}{s+2} \end{bmatrix}$$

4.12 Find a realization for each column of $\hat{\mathbf{G}}(s)$ in Problem 4.11 and then connect them, as shown in Fig. 4.4(a), to obtain a realization of $\hat{\mathbf{G}}(s)$. What is the dimension of this realization? Compare this dimension with the one in Problem 4.11.

4.13 Find a realization for each row of $\hat{\mathbf{G}}(s)$ in Problem 4.11 and then connect them, as shown in Fig. 4.4(b), to obtain a realization of $\hat{\mathbf{G}}(s)$. What is the dimension of this realization? Compare this dimension with the ones in Problems 4.11 and 4.12.

4.14 Find a realization for

$$\hat{\mathbf{G}}(s) = \begin{bmatrix} \dfrac{-(12s+6)}{3s+34} & \dfrac{22s+23}{3s+34} \end{bmatrix}$$

4.15 Consider the n-dimensional state equation

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u \qquad y = \mathbf{c}\mathbf{x}$$

Let $\hat{g}(s)$ be its transfer function. Show that $\hat{g}(s)$ has m zeros or, equivalently, the numerator of $\hat{g}(s)$ has degree m if and only if

$$\mathbf{c}\mathbf{A}^i\mathbf{b} = 0 \qquad \text{for } i = 0, 1, 2, \ldots, n-m-2$$

and $\mathbf{c}\mathbf{A}^{n-m-1}\mathbf{b} \neq 0$. Or, equivalently, the difference between the degrees of the denominator and numerator of $\hat{g}(s)$ is $\alpha = n - m$ if and only if

$$\mathbf{c}\mathbf{A}^{\alpha-1}\mathbf{b} \neq 0 \quad \text{and} \quad \mathbf{c}\mathbf{A}^i\mathbf{b} = 0$$

for $i = 0, 1, 2, \ldots, \alpha - 2$.

4.16 Find fundamental matrices and state transition matrices for

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

and

$$\dot{\mathbf{x}} = \begin{bmatrix} -1 & e^{2t} \\ 0 & -1 \end{bmatrix} \mathbf{x}$$

4.17 Show $\partial \boldsymbol{\Phi}(t_0, t)/\partial t = -\boldsymbol{\Phi}(t_0, t)\mathbf{A}(t)$.

4.18 Given

$$\mathbf{A}(t) = \begin{bmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{bmatrix}$$

show

$$\det \boldsymbol{\Phi}(t, t_0) = \exp\left[\int_{t_0}^{t} (a_{11}(\tau) + a_{22}(\tau))\, d\tau\right]$$

4.19 Let

$$\boldsymbol{\Phi}(t, t_0) = \begin{bmatrix} \boldsymbol{\Phi}_{11}(t, t_0) & \boldsymbol{\Phi}_{12}(t, t_0) \\ \boldsymbol{\Phi}_{21}(t, t_0) & \boldsymbol{\Phi}_{22}(t, t_0) \end{bmatrix}$$

be the state transition matrix of

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \mathbf{A}_{11}(t) & \mathbf{A}_{12}(t) \\ \mathbf{0} & \mathbf{A}_{22}(t) \end{bmatrix} \mathbf{x}(t)$$

Show that $\boldsymbol{\Phi}_{21}(t, t_0) = \mathbf{0}$ for all t and t_0 and that $(\partial/\partial t)\boldsymbol{\Phi}_{ii}(t, t_0) = \mathbf{A}_{ii}\boldsymbol{\Phi}_{ii}(t, t_0)$, for $i = 1, 2$.

4.20 Find the state transition matrix of

$$\dot{\mathbf{x}} = \begin{bmatrix} -\sin t & 0 \\ 0 & -\cos t \end{bmatrix} \mathbf{x}$$

4.21 Verify that $\mathbf{X}(t) = e^{\mathbf{A}t}\mathbf{C}e^{\mathbf{B}t}$ is the solution of

$$\dot{\mathbf{X}} = \mathbf{AX} + \mathbf{XB} \qquad \mathbf{X}(0) = \mathbf{C}$$

4.22 Show that if $\dot{\mathbf{A}}(t) = \mathbf{A}_1\mathbf{A}(t) - \mathbf{A}(t)\mathbf{A}_1$, then

$$\mathbf{A}(t) = e^{\mathbf{A}_1 t}\mathbf{A}(0)e^{-\mathbf{A}_1 t}$$

Show also that the eigenvalues of $\mathbf{A}(t)$ are independent of t.

4.23 Find an equivalent time-invariant state equation of the equation in Problem 4.20.

4.24 Transform a time-invariant $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ into $(\mathbf{0}, \bar{\mathbf{B}}(t), \bar{\mathbf{C}}(t))$ by a time-varying equivalence transformation.

4.25 Find a time-varying realization and a time-invariant realization of the impulse response $g(t) = t^2 e^{\lambda t}$.

4.26 Find a realization of $g(t, \tau) = \sin t (e^{-(t-\tau)}) \cos \tau$. Is it possible to find a time-invariant state equation realization?