ECE 602 Fall 2010

State Equation Solution

by

Stanislaw H. Żak

September 14, 2010

Linear systems can be modeled using convolutions and in the case of lumped systems, we can use state-space equations. Analytical computation of convolutions can be quite involved. An efficient method of obtaining a solution to a model of a lumped system is to convert the transfer function into a state-space model and then solve the state equation. We begin with a time-invariant case and then proceed with the time-varying case.

1 State Equation Solution of Uncontrolled System

Given a dynamical system's time-invariant model of the form

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) \tag{1}$$

subject to an initial condition

$$\boldsymbol{x}(0) = \boldsymbol{x}_0, \tag{2}$$

where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. The solution to (1) subject to (2) is

$$\boxed{\boldsymbol{x}(t) = e^{\boldsymbol{A}t}\boldsymbol{x}_0} \tag{3}$$

The above is a generalization of the situation when $\boldsymbol{A}$ is a scalar. The exponential matrix $e^{\boldsymbol{A}t}$ is defined as

$$e^{\boldsymbol{A}t} = \boldsymbol{I}_n + t\boldsymbol{A} + \frac{t^2}{2!}\boldsymbol{A}^2 + \frac{t^3}{3!}\boldsymbol{A}^3 + \cdots$$

To verify that (3) solves (1) subject to (2), we use the fact that

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

Suppose now that the initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_0$ is replaced with a more general condition

$$\boldsymbol{x}(t_0) = \boldsymbol{x}_0,$$

where t_0 is a given initial time instant. Then, the solution to $\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t)$ subject to $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$ is

$$\boxed{\boldsymbol{x}(t) = e^{\boldsymbol{A}(t-t_0)}\boldsymbol{x}_0}$$

The matrix $e^{\boldsymbol{A}(t-t_0)}$ is often written as

$$e^{\boldsymbol{A}(t-t_0)} = \boldsymbol{\Phi}\left(t, t_0\right)$$

and is called the *state transition matrix*, because it relates the state at any instant of time t_0 to the state at any other time t.

2 Computing State Transition Matrix of LTI Systems

For an overview of methods to calculate the state transition matrix for a linear, time-invariant (LTI) system, we refer the reader to the classic paper on the subject by Moler and Van Loan [2], where the numerical issues related to the computation of the matrix exponential are discussed. In what follows, we present a few simple methods to compute the matrix exponential.

2.1 The Inverse Laplace Transform Method

In the case when the matrix $\boldsymbol{A}$ is of low dimensions, we can use the formula that results from applying the Laplace transform to the time-invariant system $\dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x}, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0$. The Laplace transform of the above matrix differential equation is

$$s\boldsymbol{X}(s) - \boldsymbol{x}_0 = \boldsymbol{A}\boldsymbol{X}(s).$$

Hence

$$\boldsymbol{X}(s) = \left(s\boldsymbol{I}_n - \boldsymbol{A}\right)^{-1}\boldsymbol{x}_0,$$

and the inverse transform yields

$$\boxed{e^{\boldsymbol{A}t} = \mathcal{L}^{-1}\left\{\left(s\boldsymbol{I}_n - \boldsymbol{A}\right)^{-1}\right\}} \tag{4}$$

where $\mathcal{L}^{-1}$ denotes the inverse Laplace transform operator.

Example 1 We use the inverse Laplace method to compute $e^{\boldsymbol{A}t}$ for

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

We have

$$\begin{aligned} e^{\boldsymbol{A}t} &= \mathcal{L}^{-1}\left([s\boldsymbol{I}_2 - \boldsymbol{A}]^{-1}\right) \\ &= \mathcal{L}^{-1}\left(\begin{bmatrix} s & -3 \\ -2 & s-1 \end{bmatrix}^{-1}\right) \\ &= \mathcal{L}^{-1}\left(\frac{1}{(s-3)(s+2)}\begin{bmatrix} s-1 & 3 \\ 2 & s \end{bmatrix}\right) \\ &= \begin{bmatrix} 0.4e^{3t} + 0.6e^{-2t} & 0.6\left(e^{3t} - e^{-2t}\right) \\ 0.4\left(e^{3t} - e^{-2t}\right) & 0.6e^{3t} + 0.4e^{-2t} \end{bmatrix}. \end{aligned} \tag{5}$$

2.2 Diagonalization Method

A convenient way to calculate $e^{\boldsymbol{A}t}$ is when the matrix $\boldsymbol{A}$ has distinct eigenvalues. In such a case the n-by-n matrix $\boldsymbol{A}$ has n linearly independent eigenvectors as per the following theorem.

Theorem 1 *Suppose the characteristic equation $\Delta(\lambda) = 0$ has n distinct roots, $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then, there exist n linearly independent vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ such that*

$$\boldsymbol{A}\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i, \quad i = 1, 2, \ldots, n.$$

Proof See Exercise 3.

$\square$

Let

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{bmatrix}.$$

The matrix $\boldsymbol{T}$ is invertible because the eigenvectors $\boldsymbol{v}_i$, $i = 1, 2, \ldots, n$ are linearly independent. Then, because $\boldsymbol{T}^{-1}\boldsymbol{T} = \boldsymbol{I}_n$, we have

$$\begin{aligned}
\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} &= \boldsymbol{T}^{-1}\boldsymbol{A}\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n \end{bmatrix} \\
&= \boldsymbol{T}^{-1}\begin{bmatrix} \boldsymbol{A}\boldsymbol{v}_1 & \boldsymbol{A}\boldsymbol{v}_2 & \cdots & \boldsymbol{A}\boldsymbol{v}_n \end{bmatrix} \\
&= \boldsymbol{T}^{-1}\begin{bmatrix} \lambda_1\boldsymbol{v}_1 & \lambda_2\boldsymbol{v}_2 & \cdots & \lambda_n\boldsymbol{v}_n \end{bmatrix} \\
&= \boldsymbol{T}^{-1}\boldsymbol{T}\begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} \\
&= \boldsymbol{\Lambda}. \qquad (6)
\end{aligned}$$

We use the above result to compute the transition matrix $e^{\boldsymbol{A}t}$. Recall that

$$e^{\boldsymbol{A}t} = \boldsymbol{I}_n + \boldsymbol{A}t + \frac{\boldsymbol{A}^2t^2}{2!} + \frac{\boldsymbol{A}^3t^3}{3!} + \cdots$$

Note that

$$\boldsymbol{\Lambda}^k = \begin{bmatrix} \lambda_1^k & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2^k & 0 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n^k \end{bmatrix}, \quad k \geq 1$$

Hence

$$e^{\boldsymbol{\Lambda}t} = \begin{bmatrix} e^{\lambda_1 t} & 0 & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & 0 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{\lambda_n t} \end{bmatrix}.$$

From (6), we obtain

$$\boldsymbol{A} = \boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1}.$$

Therefore,

$$\begin{aligned}
\boldsymbol{A}^k &= \left(\boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1}\right)^k \\
&= \left(\boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1}\right)\left(\boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1}\right) \times \cdots \times \left(\boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1}\right) \\
&= \boldsymbol{T}\boldsymbol{\Lambda}^k\boldsymbol{T}^{-1},
\end{aligned}$$

and hence

$$
\begin{aligned}
e^{\boldsymbol{A}t} &= e^{\boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1}} \\
&= \boldsymbol{I}_n + \boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1}t + \frac{\left(\boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1}\right)^2 t^2}{2!} + \frac{\left(\boldsymbol{T}\boldsymbol{\Lambda}\boldsymbol{T}^{-1}\right)^3 t^3}{3!} + \cdots \\
&= \boldsymbol{T}\left(\boldsymbol{I}_n + \boldsymbol{\Lambda}t + \frac{\boldsymbol{\Lambda}^2 t^2}{2!} + \frac{\boldsymbol{\Lambda}^3 t^3}{3!}\right)\boldsymbol{T}^{-1} \\
&= \boldsymbol{T}e^{\boldsymbol{\Lambda}t}\boldsymbol{T}^{-1}.
\end{aligned}
\tag{7}
$$

Example 2 We use the diagonalization method to compute $e^{\boldsymbol{A}t}$ for

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

Note that this is the same matrix $\boldsymbol{A}$ as in Example 1. The eigenvalues of $\boldsymbol{A}$ are distinct so the diagonalization method is applicable here. We first compute an eigenvector, $\boldsymbol{v}_1$, corresponding to the eigenvalue $\lambda_1 = 3$ by solving the eigenvalue-eigenvector equation,

$$
\left(\boldsymbol{A} - \lambda_1 \boldsymbol{I}_2\right)\boldsymbol{v}_1 = \boldsymbol{0}.
$$

That is,

$$
\left(\boldsymbol{A} - \lambda_1 \boldsymbol{I}_2\right)\boldsymbol{v}_1 = \begin{bmatrix} -3 & 3 \\ 2 & -2 \end{bmatrix}\begin{bmatrix} v_{11} \\ v_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.
$$

We take

$$
\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.
$$

We next compute an eigenvector corresponding to the eigenvalue $\lambda_2 = -2$ by solving the equation

$$
\left(\boldsymbol{A} - \lambda_2 \boldsymbol{I}_2\right)\boldsymbol{v}_2 = \begin{bmatrix} 2 & 3 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} v_{12} \\ v_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.
$$

We take

$$
\boldsymbol{v}_2 = \begin{bmatrix} -3 \\ 2 \end{bmatrix}.
$$

We use the above eigenvectors to form the transformation matrix

$$
\boldsymbol{T} = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{bmatrix} = \begin{bmatrix} 1 & -3 \\ 1 & 2 \end{bmatrix}.
$$

Its inverse is

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

We are now ready to compute $e^{\boldsymbol{A}t}$ using (7),

$$\begin{aligned} e^{\boldsymbol{A}t} &= \boldsymbol{T}e^{\Lambda t}\boldsymbol{T}^{-1} \\ &= \frac{1}{5}\begin{bmatrix} 1 & -3 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} e^{3t} & 0 \\ 0 & e^{-2t} \end{bmatrix}\begin{bmatrix} 2 & 3 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{5}\begin{bmatrix} 2e^{3t}+3e^{-2t} & 3\left(e^{3t}-e^{-2t}\right) \\ 2\left(e^{3t}-e^{-2t}\right) & 3e^{3t}+2e^{-2t} \end{bmatrix}. \end{aligned}$$

The result is the same as when we used the inverse Laplace transform method.

The above example illustrates that the transformation $\boldsymbol{T}$ producing the diagonal form of $\boldsymbol{A}$ is a powerful computational tool. The existence of a diagonalizing $\boldsymbol{T}$ is not always guaranteed. The inability to transform $\boldsymbol{A}$ to diagonal form is connected with the existence of repeated eigenvalues that we discuss next.

3 Computing Functions of a Square Matrix

In this section, we discuss methods of computing functions of a square matrix that is not necessarily diagonalizable. The methods discussed allow us to easily compute the matrix exponential and hence the state transition matrix of LTI systems. A major tool used by the methods discussed is the so called Jordan canonical form that we discuss next.

3.1 The Jordan Canonical Form

If a square n-by-n matrix $\boldsymbol{A}$ has repeated eigenvalues, then it may not be diagonalizable. Although some matrices possessing repeated eigenvalues can be transformed into diagonal form, in many cases it is only possible to obtain the so-called "Jordan canonical form" or just the "Jordan from," for short. To illustrate the point, we consider the case when the characteristic equation, $\Delta(\lambda) = (\lambda - \lambda_1)^n$, that is, the matrix $\boldsymbol{A}$ has just one eigenvalue with multiplicity n. For simplicity, we analyze the case when $n = 4$ and suppose that

$$\operatorname{rank}\left[\begin{array}{c} \boldsymbol{A} - \lambda_1 \boldsymbol{I}_4 \end{array}\right] = 3.$$

That is, the nullity of the matrix $\left[\begin{array}{c} \boldsymbol{A} - \lambda_1 \boldsymbol{I}_4 \end{array}\right]$ is one, which means that there is only one eigenvector corresponding to the eigenvalue λ_1. Thus such a matrix $\boldsymbol{A}$ cannot be diagonalized. We need three more linearly independent vectors to represent $\boldsymbol{A}$ in a new basis, where it has a particularly simple form that is as close as possible to being diagonal.

To proceed, we need a notion of a generalized eigenvector that we define now.

Definition 1 *A vector $\boldsymbol{v}$ is called a generalized eigenvector of grade k if*

$$(\boldsymbol{A} - \lambda \boldsymbol{I}_n)^k \boldsymbol{v} = \boldsymbol{0}$$

and

$$(\boldsymbol{A} - \lambda \boldsymbol{I}_n)^{k-1} \boldsymbol{v} \neq \boldsymbol{0}.$$

◇

For $k = 1$, we have

$$(\boldsymbol{A} - \lambda \boldsymbol{I}_n) \boldsymbol{v} = \boldsymbol{0},$$

which means that $\boldsymbol{v}$ is an ordinary eigenvector corresponding to the eigenvalue λ.

Suppose now that $\boldsymbol{A} \in \mathbb{R}^{4\times 4}$ and $\boldsymbol{v}$ is a generalized eigenvector of grade $n = 4$ corresponding to the repeated eigenvalue λ_1 whose multiplicity is also 4. Then by Definition 1,

$$(\boldsymbol{A} - \lambda_1 \boldsymbol{I}_4)^4 \boldsymbol{v} = \boldsymbol{0} \quad \text{but} \quad (\boldsymbol{A} - \lambda_1 \boldsymbol{I}_4)^3 \boldsymbol{v} \neq \boldsymbol{0}.$$

Let

$$\boldsymbol{v}_1 = (\boldsymbol{A} - \lambda_1 \boldsymbol{I}_4)^3 \boldsymbol{v}.$$

Then,

$$\begin{aligned} (\boldsymbol{A} - \lambda_1 \boldsymbol{I}_4) \boldsymbol{v}_1 &= (\boldsymbol{A} - \lambda_1 \boldsymbol{I}_4)^4 \boldsymbol{v} \\ &= \boldsymbol{0}. \end{aligned}$$

Hence

$$\boldsymbol{A}\boldsymbol{v}_1 = \lambda_1 \boldsymbol{v}_1. \tag{8}$$

We also have to have

$$\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}_4\right)^2\boldsymbol{v}\neq\boldsymbol{0},$$

otherwise we would have $\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}_4\right)^3\boldsymbol{v}=\boldsymbol{0}$ as well, which would mean that $\boldsymbol{v}$ is not a generalized eigenvector of grade 4.

Let

$$\boldsymbol{v}_2=\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}_4\right)^2\boldsymbol{v}.$$

Note that

$$\begin{aligned}\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}_4\right)\boldsymbol{v}_2 &= \left(\boldsymbol{A}-\lambda_1\boldsymbol{I}_4\right)^3\boldsymbol{v}\\ &= \boldsymbol{v}_1.\end{aligned}$$

Therefore,

$$\boldsymbol{A}\boldsymbol{v}_2=\boldsymbol{v}_1+\lambda_1\boldsymbol{v}_2. \tag{9}$$

We proceed in a similar way by letting

$$\boldsymbol{v}_3=\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}_4\right)\boldsymbol{v}.$$

We observer that

$$\begin{aligned}\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}_4\right)\boldsymbol{v}_3 &= \left(\boldsymbol{A}-\lambda_1\boldsymbol{I}_4\right)^2\boldsymbol{v}\\ &= \boldsymbol{v}_2.\end{aligned}$$

Hence,

$$\boldsymbol{A}\boldsymbol{v}_3=\boldsymbol{v}_2+\lambda_1\boldsymbol{v}_3. \tag{10}$$

Finally, we let

$$\boldsymbol{v}_4=\boldsymbol{v}$$

and we note that

$$\begin{aligned}\left(\boldsymbol{A}-\lambda_1\boldsymbol{I}_4\right)\boldsymbol{v}_4 &= \left(\boldsymbol{A}-\lambda_1\boldsymbol{I}_4\right)\boldsymbol{v}\\ &= \boldsymbol{v}_3.\end{aligned}$$

Therefore,

$$\boldsymbol{A}\boldsymbol{v}_4=\boldsymbol{v}_3+\lambda_1\boldsymbol{v}_4. \tag{11}$$

Combining now (8)–(11), we obtain

$$\begin{aligned} \boldsymbol{A}\begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 & \boldsymbol{v}_4 \end{bmatrix} &= \begin{bmatrix} \lambda_1\boldsymbol{v}_1 & \boldsymbol{v}_1+\lambda_1\boldsymbol{v}_2 & \boldsymbol{v}_2+\lambda_1\boldsymbol{v}_3 & \boldsymbol{v}_3+\lambda_1\boldsymbol{v}_4 \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 & \boldsymbol{v}_4 \end{bmatrix}\begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix}. \end{aligned} \tag{12}$$

Let

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 & \boldsymbol{v}_4 \end{bmatrix}.$$

The matrix $\boldsymbol{T}$ is invertible by construction, that is, the set of vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$ forms a basis for $\mathbb{R}^4$. Therefore we can represent (12) as

$$\begin{aligned} \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} &= \begin{bmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 1 \\ 0 & 0 & 0 & \lambda_1 \end{bmatrix} \\ &= \boldsymbol{J}. \end{aligned} \tag{13}$$

The representation of $\boldsymbol{A}$ in the basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$ has the form of a Jordan block of order 4.

Example 3 Consider the matrix

$$\boldsymbol{A} = \begin{bmatrix} 5 & 2 & 4 \\ 0 & 5 & 3 \\ 0 & 0 & 5 \end{bmatrix}.$$

The characteristic polynomial is $\Delta(\lambda) = (\lambda - 5)^3$ and

$$\operatorname{rank}\begin{bmatrix} \boldsymbol{A} - \lambda\boldsymbol{I}_3 \end{bmatrix}\Big|_{\lambda=5} = 2,$$

that is, the nullity of the characteristic matrix is 1. We proceed to find $\boldsymbol{v}_1$ such that

$$\begin{aligned} \left[\boldsymbol{A} - \lambda\boldsymbol{I}_3\right]\big|_{\lambda=5}\boldsymbol{v}_1 &= \begin{bmatrix} 0 & 2 & 4 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} v_{11} \\ v_{21} \\ v_{31} \end{bmatrix} \\ &= \mathbf{0}. \end{aligned}$$

We take

$$\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

Then we proceed to find $\boldsymbol{v}_2$ such that

$$[\boldsymbol{A} - \lambda \boldsymbol{I}_3]|_{\lambda=5}\, \boldsymbol{v}_2 = \boldsymbol{v}_1,$$

that is, we solve for $\boldsymbol{v}_2$ satisfying

$$\begin{bmatrix} 0 & 2 & 4 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_{12} \\ v_{22} \\ v_{32} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}.$$

We can take

$$\boldsymbol{v}_2 = \begin{bmatrix} 0 \\ 1/2 \\ 0 \end{bmatrix}.$$

Finally we find $\boldsymbol{v}_3$ such that

$$[\boldsymbol{A} - \lambda \boldsymbol{I}_3]|_{\lambda=5}\, \boldsymbol{v}_3 = \boldsymbol{v}_2,$$

that is, we solve for $\boldsymbol{v}_3$ satisfying

$$\begin{bmatrix} 0 & 2 & 4 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_{13} \\ v_{23} \\ v_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ 1/2 \\ 0 \end{bmatrix}.$$

We can take

$$\boldsymbol{v}_3 = \begin{bmatrix} 0 \\ -2/6 \\ 1/6 \end{bmatrix}.$$

The transformation matrix $\boldsymbol{T}$ has the form

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -2/6 \\ 0 & 0 & 1/6 \end{bmatrix},$$

and

$$\boldsymbol{T}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 0 & 6 \end{bmatrix}.$$

The matrix $\boldsymbol{A}$ in the new basis has the form

$$\begin{aligned}\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 5 & 2 & 4 \\ 0 & 5 & 3 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/2 & -2/6 \\ 0 & 0 & 1/6 \end{bmatrix} \\ &= \begin{bmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix},\end{aligned}$$

which is a third order Jordan block.

If the rank of the characteristic matrix evaluated at λ_1 is $n-2$, then the equation

$$[\boldsymbol{A} - \lambda_1 \boldsymbol{I}_n]\,\boldsymbol{v} = \boldsymbol{0},$$

has two linearly independent solutions. Therefore, $\boldsymbol{A}$ has two linearly independent eigenvectors. Let $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$ be linearly independent eigenvectors associated with the eigenvalue λ_i. In order to transform $\boldsymbol{A}$ into its Jordan form, we need to find $n-2$ generalized eigenvectors. In this case we have two chains of generalized eigenvectors, $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$ and $\{\boldsymbol{v}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{v}_l\}$, where $k + l = n$. In the basis

$$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k, \boldsymbol{v}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{v}_l\}$$

the matrix $\boldsymbol{A}$ will have a block diagonal form

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{bmatrix} \boldsymbol{J}_1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{J}_2 \end{bmatrix},$$

where $\boldsymbol{J}_1$ is a Jordan block of order k and $\boldsymbol{J}_2$ is a Jordan block of order l.

Example 4 Let

$$\boldsymbol{A} = \begin{bmatrix} 3 & 1 & -1 \\ 0 & 6 & -3 \\ 0 & 3 & 0 \end{bmatrix}.$$

The characteristic polynomial

$$\Delta(\lambda) = (\lambda - 3)^3$$

and

$$\begin{aligned} \operatorname{rank}\left[\boldsymbol{A}-\lambda \boldsymbol{I}_3\right]\big|_{\lambda=3} &= \begin{bmatrix} 0 & 1 & -1 \\ 0 & 3 & -3 \\ 0 & 3 & -3 \end{bmatrix} \\ &= 1 \end{aligned}$$

As two linearly independent solutions to

$$\left[\boldsymbol{A}-\lambda \boldsymbol{I}_3\right]\big|_{\lambda=3} \boldsymbol{v}=\mathbf{0},$$

we can take

$$\boldsymbol{u}_1=\begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix} \quad \text{and} \quad \boldsymbol{v}_1=\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}.$$

Now we have to find one generalized eigenvector that together with $\boldsymbol{u}_1$ and $\boldsymbol{v}_1$ would form a basis for $\mathbb{R}^3$ and enable us to construct a transformation that brings $\boldsymbol{A}$ into its Jordan form. We try first $\boldsymbol{u}_1$ to construct $\boldsymbol{u}_2$ such that

$$\left[\boldsymbol{A}-\lambda \boldsymbol{I}_3\right]\big|_{\lambda=3} \boldsymbol{u}_2=\boldsymbol{u}_1,$$

that is, we solve for $\boldsymbol{u}_2$ satisfying

$$\begin{bmatrix} 0 & 1 & -1 \\ 0 & 3 & -3 \\ 0 & 3 & -3 \end{bmatrix}\begin{bmatrix} u_{12} \\ u_{22} \\ u_{32} \end{bmatrix}=\begin{bmatrix} 1 \\ 3 \\ 3 \end{bmatrix}.$$

We can take

$$\boldsymbol{u}_2=\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

We form the transformation matrix

$$\boldsymbol{T}=\begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{v}_1 \end{bmatrix}=\begin{bmatrix} 1 & 1 & 0 \\ 3 & 1 & 1 \\ 3 & 0 & 1 \end{bmatrix}.$$

Its inverse is

$$\boldsymbol{T}^{-1}=\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ -3 & 3 & -2 \end{bmatrix}.$$

It is then easy to check that

$$\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{bmatrix} \boldsymbol{J}_1 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{J}_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

where

$$\boldsymbol{J}_1 = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} \quad \text{and} \quad \boldsymbol{J}_2 = 3.$$

In general, the Jordan canonical form is a block diagonal matrix, where each diagonal block is a Jordan block corresponding to a possibly different eigenvalue, that is,

$$\begin{aligned} \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} &= \begin{bmatrix} \begin{matrix} \lambda_1 & 1 & 0 & \cdots & 0 \\ 0 & \lambda_1 & 1 & \cdots & 0 \\ 0 & 0 & \lambda_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_1 \end{matrix} & & \\ & \ddots & \\ & & \begin{matrix} \lambda_r & 1 & 0 & \cdots & 0 \\ 0 & \lambda_r & 1 & \cdots & 0 \\ 0 & 0 & \lambda_r & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_r \end{matrix} \end{bmatrix} \\ &= \begin{bmatrix} \boldsymbol{J}_1 & \boldsymbol{O} & \cdots & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{J}_2 & & \boldsymbol{O} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{J}_q \end{bmatrix}, \end{aligned} \tag{14}$$

where each $r_i \times r_i$ Jordan block has the form $\boldsymbol{J}_i = \lambda_i$ if $r_i = 1$, or, for $r_i > 1$ the corresponding Jordan block has a non-diagonal structure. Note that

$$n = r_1 + r_2 + \cdots + r_q$$

and

$$q \geq r,$$

which means that there may be more than one Jordan blocks associated with one eigenvalue as was the case in Example 4.

Example 5 To illustrate the general case depicted in (14), we consider, as in Chen [1, p. 60], a 5-by-5 matrix $\boldsymbol{A}$ with repeated eigenvalue λ_1 of multiplicity 4 and a simple eigenvalue λ_2, that is, the characteristic polynomial of $\boldsymbol{A}$ is

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

Depending on the nullity of the matrix $[\boldsymbol{A} - \lambda_1 \boldsymbol{I}_5]$, the Jordan form of the matrix $\boldsymbol{A}$ make take one of the following forms:

- If the nullity of $[\boldsymbol{A} - \lambda_1 \boldsymbol{I}_5]$ is 1, then there is only one eigenvector, say $\boldsymbol{u}_1$ associated with λ_1 and we need to compute three generalized eigenvectors to form a basis $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3, \boldsymbol{u}_4, \boldsymbol{v}_1\}$, where $\boldsymbol{v}_1$ is an eigenvector associated with λ_2. The matrix $\boldsymbol{A}$ in this basis will have the form:

$$\hat{\boldsymbol{A}}_1 = \boldsymbol{T}_1^{-1}\boldsymbol{A}\boldsymbol{T}_1 = \begin{bmatrix} \lambda_1 & 1 & 0 & 0 & \vdots & 0 \\ 0 & \lambda_1 & 1 & 0 & \vdots & 0 \\ 0 & 0 & \lambda_1 & 1 & \vdots & 0 \\ 0 & 0 & 0 & \lambda_1 & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \vdots & \lambda_2 \end{bmatrix}.$$

  Note that we can re-order the basis vectors as $\{\boldsymbol{v}_1, \boldsymbol{u}_1, \boldsymbol{u}_1, \boldsymbol{u}_3, \boldsymbol{u}_4\}$. In this basis the matrix $\boldsymbol{A}$ representation has the form

$$\hat{\boldsymbol{A}}_2 = \boldsymbol{T}_2^{-1}\boldsymbol{A}\boldsymbol{T}_2 = \begin{bmatrix} \lambda_2 & \vdots & 0 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \vdots & \lambda_1 & 1 & 0 & 0 \\ 0 & \vdots & 0 & \lambda_1 & 1 & 0 \\ 0 & \vdots & 0 & 0 & \lambda_1 & 1 \\ 0 & \vdots & 0 & 0 & 0 & \lambda_1 \end{bmatrix}.$$

- If the nullity of $[\boldsymbol{A} - \lambda_1 \boldsymbol{I}_5]$ is 2, then there are two eigenvectors associated with λ_1, that is, there are two Jordan blocks associated with λ_1. The matrix $\boldsymbol{A}$ may take one of the following forms:

$$\hat{\boldsymbol{A}}_3 = \boldsymbol{T}_3^{-1}\boldsymbol{A}\boldsymbol{T}_3 = \begin{bmatrix} \lambda_1 & 1 & 0 & \vdots & 0 & & 0 \\ 0 & \lambda_1 & 1 & \vdots & 0 & & 0 \\ 0 & 0 & \lambda_1 & \vdots & 0 & & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & 0 & \vdots & \lambda_1 & \vdots & 0 \\ & & & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & & 0 & \vdots & \lambda_2 \end{bmatrix}$$

or

$$\hat{\boldsymbol{A}}_4 = \boldsymbol{T}_4^{-1}\boldsymbol{A}\boldsymbol{T}_4 = \begin{bmatrix} \lambda_1 & 1 & \vdots & 0 & 0 & & 0 \\ 0 & \lambda_1 & \vdots & 0 & 0 & & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & \vdots & \lambda_1 & 1 & \vdots & 0 \\ 0 & 0 & \vdots & 0 & \lambda_1 & \vdots & 0 \\ & & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & & 0 & 0 & \vdots & \lambda_2 \end{bmatrix}.$$

- If the nullity of $[\boldsymbol{A} - \lambda_1 \boldsymbol{I}_5]$ is 3, then there are three eigenvectors associated with λ_1, that is, there are three Jordan blocks associated with λ_1. The matrix $\boldsymbol{A}$ may take the following form:

$$\hat{\boldsymbol{A}}_5 = \boldsymbol{T}_5^{-1}\boldsymbol{A}\boldsymbol{T}_5 = \begin{bmatrix} \lambda_1 & 1 & \vdots & 0 & & 0 & & 0 \\ 0 & \lambda_1 & \vdots & 0 & & 0 & & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & & & \\ 0 & 0 & \vdots & \lambda_1 & \vdots & 0 & & 0 \\ & & \cdots & \cdots & \cdots & \cdots & \cdots & \\ 0 & 0 & & 0 & \vdots & \lambda_1 & \vdots & 0 \\ & & & & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & & 0 & & 0 & \vdots & \lambda_2 \end{bmatrix}.$$

- Finally, if the nullity of $[\boldsymbol{A} - \lambda_1 \boldsymbol{I}_5]$ is 4, then there are four eigenvectors associated with λ_1, that is, there are four Jordan blocks associated with λ_1. The Jordan form of the matrix matrix $\boldsymbol{A}$ is a diagonal matrix,

$$\hat{\boldsymbol{A}}_6 = \boldsymbol{T}_6^{-1}\boldsymbol{A}\boldsymbol{T}_6 = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{bmatrix}.$$

To summarize our discussion, we need the following definitions.

Definition 2 *For a given eigenvalue λ_i of an n-by-n matrix $\boldsymbol{A}$, its algebraic multiplicity is the multiplicity of it in the characteristics polynomial $\Delta(\lambda)$, that is, the algebraic multiplicity of λ_i is the sum of orders of all Jordan blocks associated with λ_i.*

◇

Definition 3 *For a given eigenvalue λ_i of an n-by-n matrix $\boldsymbol{A}$, its geometric multiplicity is the nullity of the associated characteristic matrix $\left[\ \boldsymbol{A} - \lambda_i \boldsymbol{I}_n\ \right]$, that is, the geometric multiplicity of λ_i equals the number of Jordan blocks associated with it.*

◇

Then we have the following theorem:

Theorem 2 *A given n-by-n matrix $\boldsymbol{A}$ is diagonalizable if and only if geometric multiplicities of its eigenvalues are the same as their algebraic multiplicities, that is, if and only if $\boldsymbol{A}$ has n linearly independent eigenvectors.*

We close our discussion on the Jordan canonical form with an observation by Owens [3, p. 34] that in almost all situations of practical relevance, the matrix $\boldsymbol{A}$ has distinct eigenvalues. So in a sense the cases described in this section can be considered "pathological" from the applications point of view. In the literature, see for example [1, p. 61], if $\boldsymbol{A}$ cannot be diagonalized then it is called to be defective. Calculation of the Jordan canonical is numerically ill conditioned. One can try MATLAB's function `jordan`. However, the matrix $\boldsymbol{A}$ must be known exactly, which in MATLAB's language means that its elements must be integers or ratios of small integers. Finally, to diagonalize a square matrix using MATLAB, one can call the function `eig`. If, however, $\boldsymbol{A}$ is not diagonalizable, then the function `eig` will give incorrect solution.

3.2 Computing Polynomial Functions of a Square Matrix

If $\boldsymbol{A}$ is an n-by-n square matrix then its k-th power is

$$\boldsymbol{A}^k = \underbrace{\boldsymbol{A}\boldsymbol{A} \times \cdots \times \boldsymbol{A}}_{k \text{ times}},$$

where

$$\boldsymbol{A}^0 = \boldsymbol{I}_n.$$

Suppose now that we have a polynomial

$$f(\lambda) = \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + \alpha_1\lambda + \alpha_0.$$

Then we define $f(\boldsymbol{A})$ as

$$f(\boldsymbol{A}) = \boldsymbol{A}^n + \alpha_{n-1}\boldsymbol{A}^{n-1} + \cdots + \alpha_1\boldsymbol{A} + \alpha_0\boldsymbol{I}_n.$$

For example, if

$$f(\lambda) = \lambda^2 + 2\lambda - 3 = (\lambda + 3)(\lambda - 1),$$

then

$$f(\boldsymbol{A}) = \boldsymbol{A}^2 + 2\boldsymbol{A} - 3\boldsymbol{I}_n = (\boldsymbol{A} + 3\boldsymbol{I}_n)(\boldsymbol{A} - \boldsymbol{I}_n).$$

Now if the matrix $\boldsymbol{A}$ is block diagonal, for example, such as

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_2 & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{A}_3 \end{bmatrix},$$

where $\boldsymbol{A}_i$ are square matrices and $\boldsymbol{O}$ denotes a matrix of zeros of appropriate dimensions, then we have

$$\boldsymbol{A}^k = \begin{bmatrix} \boldsymbol{A}_1^k & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{A}_2^k & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & \boldsymbol{A}_3^k \end{bmatrix}$$

and

$$f(\boldsymbol{A}) = \begin{bmatrix} f(\boldsymbol{A}_1) & \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{O} & f(\boldsymbol{A}_2) & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & f(\boldsymbol{A}_3) \end{bmatrix}.$$

Consider a similar matrix $\hat{\boldsymbol{A}}$ of the given matrix $\boldsymbol{A}$, that is,

$$\hat{\boldsymbol{A}} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T},$$

where $\boldsymbol{T} \in \mathbb{R}^{n\times n}$ is an invertible matrix. The above can be re-written as

$$\boldsymbol{A} = \boldsymbol{T}\hat{\boldsymbol{A}}\boldsymbol{T}^{-1},$$

and we have

$$\begin{aligned} \boldsymbol{A}^k &= \left(\boldsymbol{T}\hat{\boldsymbol{A}}\boldsymbol{T}^{-1}\right)\left(\boldsymbol{T}\hat{\boldsymbol{A}}\boldsymbol{T}^{-1}\right)\times\cdots\times\left(\boldsymbol{T}\hat{\boldsymbol{A}}\boldsymbol{T}^{-1}\right) \\ &= \boldsymbol{T}\hat{\boldsymbol{A}}^k\boldsymbol{T}^{-1}, \end{aligned}$$

or

$$\begin{aligned} \hat{\boldsymbol{A}}^k &= \left(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}\right)\left(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}\right)\times\cdots\times\left(\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}\right) \\ &= \boldsymbol{T}^{-1}\boldsymbol{A}^k\boldsymbol{T}. \end{aligned}$$

Thus, we have

$$f\left(\boldsymbol{A}\right) = \boldsymbol{T}f(\hat{\boldsymbol{A}})\boldsymbol{T}^{-1} \tag{15}$$

or

$$f(\hat{\boldsymbol{A}}) = \boldsymbol{T}^{-1}f\left(\boldsymbol{A}\right)\boldsymbol{T}. \tag{16}$$

Now let λ_i be the i-th eigenvalue of $\boldsymbol{A}$ and suppose that the algebraic multiplicity of λ_i is n_i. Then the characteristic polynomial of $\boldsymbol{A}$ can be written as,

$$\Delta(\lambda) = \det(\lambda\boldsymbol{I}_n - \boldsymbol{A}) = \prod_{i=1}^{r}\left(\lambda - \lambda_i\right)^{n_i}. \tag{17}$$

We now discuss a very useful property of Jordan blocks that we will utilize in the proof of the very imporatnt theorem of linear algebra, the theorem of Cayley and Hamilton. Suppose that we have a fourth order Jordan block,

$$\boldsymbol{J} = \begin{bmatrix} \lambda_i & 1 & 0 & 0 \\ 0 & \lambda_i & 1 & 0 \\ 0 & 0 & \lambda_i & 1 \\ 0 & 0 & 0 & \lambda_i \end{bmatrix}.$$

Then we have

$$(\boldsymbol{J} - \lambda_i \boldsymbol{I}) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Let us now compute successive powers of $(\boldsymbol{J} - \lambda \boldsymbol{I})$. We obtain

$$(\boldsymbol{J} - \lambda_i \boldsymbol{I})^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (\boldsymbol{J} - \lambda_i \boldsymbol{I})^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and

$$(\boldsymbol{J} - \lambda_i \boldsymbol{I})^k = \boldsymbol{O} \quad \text{for} \quad k \geq 4.$$

A matrix with above property is called a *nilpotent* matrix.

In general, if we have a Jordan block of order r_i then

$$(\boldsymbol{J} - \lambda_i \boldsymbol{I})^k = \boldsymbol{O} \quad \text{for} \quad k \geq r_i. \tag{18}$$

We have all we need to state and prove the theorem of Cayley and Hamilton.

Theorem 3 *(Cayley-Hamilton theorem) A square matrix satisfies its characteristic polynomial equation, that is,*

$$\Delta(\boldsymbol{A}) = \boldsymbol{O}.$$

Proof Suppose that $\hat{\boldsymbol{A}}$ is the Jordan canonical form of $\boldsymbol{A}$. Thus there exists a non-singular matrix $\boldsymbol{T}$ such that

$$\hat{\boldsymbol{A}} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = \begin{bmatrix} \boldsymbol{J}_1 & \boldsymbol{O} & \cdots & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{J}_2 & & \boldsymbol{O} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{J}_q \end{bmatrix}.$$

Let us now consider the matrix polynomial $\Delta(\boldsymbol{A})$. We use (15) to obtain

$$
\begin{aligned}
\Delta(\boldsymbol{A}) &= \boldsymbol{T}\Delta(\hat{\boldsymbol{A}})\boldsymbol{T}^{-1} \\
&= \boldsymbol{T}\prod_{i=1}^{r}\left(\hat{\boldsymbol{A}}-\lambda_i\boldsymbol{I}_n\right)^{n_i}\boldsymbol{T}^{-1} \\
&= \boldsymbol{T}\prod_{i=1}^{r}\begin{bmatrix} \boldsymbol{J}_1-\lambda_i\boldsymbol{I}_{r_1} & \boldsymbol{O} & \cdots & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{J}_2-\lambda_i\boldsymbol{I}_{r_2} & & \boldsymbol{O} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{J}_q-\lambda_i\boldsymbol{I}_{r_q} \end{bmatrix}^{n_i}\boldsymbol{T}^{-1} \\
&= \boldsymbol{T}\prod_{i=1}^{r}\begin{bmatrix} (\boldsymbol{J}_1-\lambda_i\boldsymbol{I}_{r_1})^{n_i} & \boldsymbol{O} & \cdots & \boldsymbol{O} \\ \boldsymbol{O} & (\boldsymbol{J}_2-\lambda_i\boldsymbol{I}_{r_2})^{n_i} & & \boldsymbol{O} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & (\boldsymbol{J}_q-\lambda_i\boldsymbol{I}_{r_q})^{n_i} \end{bmatrix}\boldsymbol{T}^{-1}.
\end{aligned}
$$

The order of any Jordan block associated with λ_i cannot exceed the corresponding algebraic multiplicity. Taking into account the nilpotent property given by (18), we obtain

$$
\begin{aligned}
\Delta(\boldsymbol{A}) &= \boldsymbol{T}\begin{bmatrix} \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{O} \\ \boldsymbol{O} & (\boldsymbol{J}_2-\lambda_1\boldsymbol{I}_{r_2})^{n_1} & & \boldsymbol{O} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & (\boldsymbol{J}_q-\lambda_1\boldsymbol{I}_{r_q})^{n_1} \end{bmatrix}\times \\
&\quad\cdots\times\begin{bmatrix} (\boldsymbol{J}_1-\lambda_r\boldsymbol{I}_{r_1})^{n_r} & \boldsymbol{O} & \cdots & \boldsymbol{O} \\ \boldsymbol{O} & (\boldsymbol{J}_2-\lambda_r\boldsymbol{I}_{r_2})^{n_r} & & \boldsymbol{O} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{O} \end{bmatrix}\boldsymbol{T}^{-1} \\
&= \boldsymbol{T}\begin{bmatrix} \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{O} & & \boldsymbol{O} \\ \vdots & & \ddots & \vdots \\ \boldsymbol{O} & \boldsymbol{O} & \cdots & \boldsymbol{O} \end{bmatrix}\boldsymbol{T}^{-1} \\
&= \boldsymbol{O}, \qquad (19)
\end{aligned}
$$

which completes the proof of the theorem.

□

It follows from the Cayley-Hamilton theorem that

$$\Delta(\boldsymbol{A}) = \boldsymbol{A}^n + \alpha_{n-1}\boldsymbol{A}^{n-1} + \cdots + \alpha_1\boldsymbol{A} + \alpha_0\boldsymbol{I}_n = \boldsymbol{O},$$

that is,

$$\boldsymbol{A}^n = -\alpha_{n-1}\boldsymbol{A}^{n-1} - \cdots - \alpha_1\boldsymbol{A} - \alpha_0\boldsymbol{I}_n. \tag{20}$$

Thus $\boldsymbol{A}^n$ can be expressed as a linear combination of the matrices

$$\{\boldsymbol{I}_n, \boldsymbol{A}, \ldots, \boldsymbol{A}^{n-1}\}.$$

If we multiply both sides of (20) by $\boldsymbol{A}$, then we obtain

$$\boldsymbol{A}^{n+1} = -\alpha_{n-1}\boldsymbol{A}^{n} - \cdots - \alpha_1\boldsymbol{A}^2 - \alpha_0\boldsymbol{A}. \tag{21}$$

Thus, $\boldsymbol{A}^{n+1}$ can be expressed as a linear combination of

$$\{\boldsymbol{A}, \boldsymbol{A}^2, \ldots, \boldsymbol{A}^{n}\}.$$

Using (20), we conclude that $\boldsymbol{A}^{n+1}$ can be expressed as a linear combination of $\{\boldsymbol{I}_n, \boldsymbol{A}, \ldots, \boldsymbol{A}^{n-1}\}$. Using the same argument, we conclude that any $\boldsymbol{A}^k$, $k \geq 0$, can be expressed as a linear combination of $\{\boldsymbol{I}_n, \boldsymbol{A}, \ldots, \boldsymbol{A}^{n-1}\}$. Therefore, any matrix polynomial $f(\boldsymbol{A})$ can also be expressed as

$$f(\boldsymbol{A}) = \beta_0\boldsymbol{I}_n + \beta_1\boldsymbol{A} + \cdots + \beta_{n-1}\boldsymbol{A}^{n-1}$$

for some β_i, $i = 0, 1, \ldots, n-1$. The above property is used to efficiently compute matrix polynomials of $\boldsymbol{A}$, which can then be used to compute the matrix exponential $e^{\boldsymbol{A}t}$ as well as the state transition matrix for discrete-time systems. In what follows, we present such a method and illustrate it with examples. We follow the presentation of Chen [1, pp. 63–67].

Suppose we are given a polynomial $f(\lambda)$ and we are interested in calculating $f(\boldsymbol{A})$. We begin with dividing $f(\lambda)$ by $\Delta(\lambda)$ to obtain

$$f(\lambda) = q(\lambda)\Delta(\lambda) + h(\lambda), \tag{22}$$

where $q(\lambda)$ is the quotient, $h(\lambda)$ is the remainder, and

$$\deg_\lambda h(\lambda) < n.$$

We now evaluate $f(\boldsymbol{A})$ using (22) and the Cayley-Hamilton theorem to obtain

$$\begin{aligned} f(\boldsymbol{A}) &= q(\boldsymbol{A})\Delta(\boldsymbol{A}) + h(\boldsymbol{A}) \\ &= q(\boldsymbol{A})\boldsymbol{O} + h(\boldsymbol{A}) \\ &= h(\boldsymbol{A}). \end{aligned}$$

Thus instead of evaluating the matrix polynomial $f(\boldsymbol{A})$ that may contain very high powers of $\boldsymbol{A}$, we can find $f(\boldsymbol{A})$ by computing $h(\boldsymbol{A})$, where the highest power of $\boldsymbol{A}$ is less than n. In addition, rather than performing the division (22), we can find $h(\lambda)$ by assuming that it has the form

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

and then solve n equations for β_i, $i = 0, 1, \ldots, n-1$. Consider first the case when $\boldsymbol{A}$ has distinct eigenvalues. To obtain equations involving the coefficients β_i, we note that

$$f(\lambda_i) = q(\lambda_i)\Delta(\lambda_i) + h(\lambda_i) = h(\lambda_i), \quad i = 1, 2, \ldots, n.$$

Example 6 For a given

$$f(\lambda) = \lambda^7,$$

we compute

$$f(\boldsymbol{A}) = \boldsymbol{A}^7,$$

where

$$\boldsymbol{A} = \begin{bmatrix} 0 & 2 \\ 1 & 1 \end{bmatrix}. \tag{23}$$

The characteristic polynomial of $\boldsymbol{A}$ is

$$\Delta(\lambda) = \lambda^2 - \lambda - 2 = (\lambda + 1)(\lambda - 2).$$

Let

$$h(\lambda) = \beta_0 + \beta_1\lambda.$$

Then we have two equations:

$$f(\lambda_i) = h(\lambda_i), \quad i = 1, 2,$$

that is,

$$f(-1) = (-1)^7 = -1 = \beta_0 - \beta_1$$

and

$$f(2) = 2^7 = 128 = \beta_0 + 2\beta_1.$$

We represent the above linear equations for β_0 and β_1 in the matrix format as

$$\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} -1 \\ 128 \end{bmatrix}.$$

Solving, we obtain $\beta_0 = 42$ and $\beta_1 = 43$. Hence

$$f(\boldsymbol{A}) = \boldsymbol{A}^7 = 42\boldsymbol{I}_2 + 43\boldsymbol{A} = \begin{bmatrix} 42 & 86 \\ 43 & 85 \end{bmatrix}.$$

Example 7 We can use the above method to evaluate $e^{\boldsymbol{A}t}$, where $\boldsymbol{A}$ is given by (23). In other words, let $f(\lambda) = e^{\lambda t}$; we wish to find

$$f(\boldsymbol{A}) = e^{\boldsymbol{A}t}.$$

Let $h(\lambda) = h(\lambda) = \beta_0 + \beta_1\lambda$. Then

$$f(-1) = h(-1)$$

yields

$$e^{-t} = \beta_0 - \beta_1,$$

and

$$f(2) = h(2)$$

yields

$$e^{2t} = \beta_0 + 2\beta_1.$$

We thus have two equations for β_0 and β_1,

$$\begin{bmatrix} 1 & -1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} e^{-t} \\ e^{2t} \end{bmatrix}.$$

Solving the above gives

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} e^{-t} \\ e^{2t} \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 2e^{-t} + e^{2t} \\ -e^{-t} + e^{2t} \end{bmatrix}.$$

Hence

$$e^{\boldsymbol{A}t} = \beta_0\boldsymbol{I}_2 + \beta_1\boldsymbol{A} = \frac{1}{3}\begin{bmatrix} 2e^{-t} + e^{2t} & -2e^{-t} + 2e^{2t} \\ -e^{-t} + e^{2t} & e^{-t} + 2e^{2t} \end{bmatrix}.$$

When the characteristic polynomial has repeated zeros, we use the derivatives of $f(\lambda)$ and $h(\lambda)$ to obtain additional equations to determine the coefficients of $h(\lambda)$ as per the following theorem.

Theorem 4 *Chen [1, p. 64] We are given $f(\lambda)$ and an $n \times n$ matrix $\boldsymbol{A}$ with characteristic polynomial*

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

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

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

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:

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

where

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

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

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

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

Example 8 We compute $e^{\boldsymbol{A}t}$ for

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

The characteristic polynomial is

$$\Delta(\lambda) = (\lambda - 1)^2 (\lambda - 2).$$

Let $f(\lambda) = e^{\lambda t}$ and let $h(\lambda) = \beta_0 + \beta_1 \lambda + \beta_2 \lambda^2$. Then using Theorem 4, we obtain

$$\begin{aligned} f(1) &= h(1), \quad e^t = \beta_0 + \beta_1 + \beta_2, \\ f'(1) &= h'(1), \quad te^t = \beta_1 + 2\beta_2, \\ f(2) &= h(2), \quad e^{2t} = \beta_0 + 2\beta_1 + 4\beta_2. \end{aligned}$$

Thus we have a set of equations

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} e^t \\ te^t \\ e^{2t} \end{bmatrix}.$$

Solving the above gives

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 0 & -2 & 1 \\ 2 & 3 & -2 \\ -1 & -1 & 1 \end{bmatrix} \begin{bmatrix} e^t \\ te^t \\ e^{2t} \end{bmatrix} = \begin{bmatrix} -2te^t + e^{2t} \\ 2e^t + 3te^t - 2e^{2t} \\ -e^t - te^t + e^{2t} \end{bmatrix}.$$

Hence

$$\begin{aligned} e^{\boldsymbol{A}t} &= h(\boldsymbol{A}) \\ &= \beta_0 \boldsymbol{I}_3 + \beta_1 \boldsymbol{A} + \beta_2 \boldsymbol{A}^2 \\ &= \left(-2te^t + e^{2t}\right) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \left(2e^t + 3te^t - 2e^{2t}\right) \begin{bmatrix} 3 & 0 & 1 \\ -1 & 1 & 2 \\ -2 & 0 & 0 \end{bmatrix} \\ &\quad + \left(-e^t - te^t + e^{2t}\right) \begin{bmatrix} 7 & 0 & 3 \\ -8 & 1 & 1 \\ -6 & 0 & -2 \end{bmatrix} \\ &= \begin{bmatrix} -e^t + 2e^{2t} & 0 & -e^t + e^{2t} \\ 6e^t + 5te^t - 6e^{2t} & e^t & 3e^t + 5te^t - 3e^{2t} \\ 2e^t - 2e^{2t} & 0 & 2e^t - e^{2t} \end{bmatrix}. \end{aligned}$$

Example 9 Suppose we are given the Jordan block of order n_i,

$$\boldsymbol{J} = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{bmatrix}.$$

The characteristic polynomial of $\boldsymbol{J}$ is $\Delta(\lambda) = (\lambda - \lambda_i)^{n_i}$. For computational simplicity we select the $(n_i - 1)$-st order polynomial $h(\lambda)$ of the form

$$h(\lambda) = \beta_0 + \beta_1(\lambda - \lambda_i) + \beta_2(\lambda - \lambda_i)^2 + \cdots + \beta_{n_i-1}(\lambda - \lambda_i)^{n_i-1}.$$

The condition $f(\lambda) = h(\lambda)$ on the spectrum of $\boldsymbol{J}$ yields

$$\beta_0 = f(\lambda_i), \quad \beta_1 = f'(\lambda_i), \quad \beta_2 = \frac{f''(\lambda_i)}{2!}, \ldots, \beta_{n_i-1} = \frac{f^{(n_i-1)}(\lambda_i)}{(n_i-1)!}.$$

Hence

$$\begin{aligned} f(\boldsymbol{J}) &= f(\lambda_i)\boldsymbol{I}_{n_i} + \frac{f'(\lambda_i)}{1!}(\boldsymbol{J} - \lambda_i\boldsymbol{I}_{n_i}) + \frac{f''(\lambda_i)}{2!}(\boldsymbol{J} - \lambda_i\boldsymbol{I}_{n_i})^2 + \cdots \\ &\qquad \cdots + \frac{f^{(n_i-1)}(\lambda_i)}{(n_i-1)!}(\boldsymbol{J} - \lambda_i\boldsymbol{I}_{n_i})^{n_i-1}. \end{aligned}$$

Using the nilpotent property of the matrix $(\boldsymbol{J} - \lambda_i\boldsymbol{I}_{n_i})$, we obtain

$$f(\boldsymbol{J}) = \begin{bmatrix} f(\lambda_i) & f'(\lambda_i)/1! & f''(\lambda_i)/2! & \cdots & f^{(n_i-1)}(\lambda_i)/(n_i-1)! \\ 0 & f(\lambda_i) & f'(\lambda_i)/1! & \cdots & f^{(n_i-2)}(\lambda_i)/(n_i-2)! \\ 0 & 0 & f(\lambda_i) & \cdots & f^{(n_i-3)}(\lambda_i)/(n_i-3)! \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ddots & f'(\lambda_i)/1! \\ 0 & 0 & 0 & \cdots & f(\lambda_i) \end{bmatrix}.$$

In particular, if $f(\lambda) = e^{\lambda t}$, then we have

$$e^{\boldsymbol{J}t} = \begin{bmatrix} e^{\lambda_i t} & te^{\lambda_i t} & t^2 e^{\lambda_i t}/2! & \cdots & t^{n_i-1}e^{\lambda_i t}/(n_i-1)! \\ 0 & e^{\lambda_i t} & te^{\lambda_i t} & \cdots & t^{n_i-2}e^{\lambda_i t}/(n_i-2)! \\ 0 & 0 & e^{\lambda_i t} & \cdots & t^{n_i-3}e^{\lambda_i t}/(n_i-3)! \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ddots & te^{\lambda_i t} \\ 0 & 0 & 0 & \cdots & e^{\lambda_i t} \end{bmatrix}.$$

4 State Equation Solution of Controlled System

Consider a model of a controlled dynamic system,

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t) \tag{24}$$

subject to a given initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_0$, where $\boldsymbol{B}$ is an $n \times m$ matrix, and $m < n$ as is usually in practice. Premultiplying both sides of (24) by the integrating factor, $e^{-\boldsymbol{A}t}$, yields

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

We then rearrange the above equation as

$$e^{-\boldsymbol{A}t}\dot{\boldsymbol{x}}(t) - e^{-\boldsymbol{A}t}\boldsymbol{A}\boldsymbol{x}(t) = e^{-\boldsymbol{A}t}\boldsymbol{B}\boldsymbol{u}(t). \tag{25}$$

Noting that

$$\frac{d}{dt}\left(e^{-\boldsymbol{A}t}\boldsymbol{x}(t)\right) = -\boldsymbol{A}e^{-\boldsymbol{A}t}\boldsymbol{x}(t) + e^{-\boldsymbol{A}t}\dot{\boldsymbol{x}}(t),$$

we can represent (25) in the form

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

Integrating both sides of the resulting equation gives

$$e^{-\boldsymbol{A}t}\boldsymbol{x}(t) - \boldsymbol{x}(0) = \int_0^t e^{-\boldsymbol{A}\tau}\boldsymbol{B}\boldsymbol{u}(\tau)d\tau.$$

Hence

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

If the initial condition is $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$, then integration from t_0 to t yields

$$\boxed{\boldsymbol{x}(t) = e^{\boldsymbol{A}(t-t_0)}\boldsymbol{x}(t_0) + \int_{t_0}^t e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)d\tau} \tag{26}$$

As in the case of the uncontrolled system, we can also use the Laplace transform technique to obtain the solution to (24) subject to the initial condition $\boldsymbol{x}(0) = \boldsymbol{x}_0$. Indeed, taking the Laplace transform of (24) and performing some manipulations yields

$$\boldsymbol{X}(s) = (sI_n - \boldsymbol{A})^{-1}\boldsymbol{x}(0) + (sI_n - \boldsymbol{A})^{-1}\boldsymbol{B}\boldsymbol{U}(s).$$

Hence

$$\boldsymbol{x}(t) = \mathcal{L}^{-1}\left\{(sI_n - \boldsymbol{A})^{-1}\boldsymbol{x}(0)\right\} + \mathcal{L}^{-1}\left\{(sI_n - \boldsymbol{A})^{-1}\boldsymbol{B}\boldsymbol{U}(s)\right\}.$$

If $\boldsymbol{x}_0 = \boldsymbol{x}(t_0)$, then

$$\boldsymbol{x}(t) = \mathcal{L}^{-1}\left\{e^{-t_0 s}(sI_n - \boldsymbol{A})^{-1}\boldsymbol{x}(t_0)\right\} + \mathcal{L}^{-1}\left\{e^{-t_0 s}(sI_n - \boldsymbol{A})^{-1}\boldsymbol{B}\boldsymbol{U}(s)\right\}.$$

5 From Continuous to Discrete Models

Our objective here is to present a method for constructing linear discrete-time models from given linear continuous-time models.

We use a sample-and-hold device that transforms a continuous signal, $f(t)$, into the staircase signal,

$$f(kh), \quad kh \leq t < (k+1)h,$$

where h is the sampling period. The sample and zero-order hold (ZOH) operation is illustrated in Figure 1.

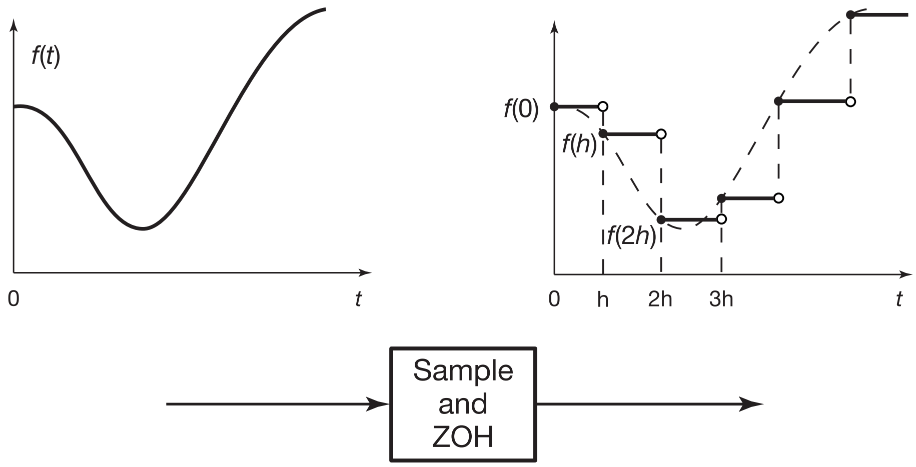

Figure 1: Sample and zero-order hold (ZOH) element operating on a continuous function.

Suppose that we are given a continuous-time model,

$$\begin{aligned} \dot{\boldsymbol{x}}(t) &= \boldsymbol{A}\boldsymbol{x}(t) + \boldsymbol{B}\boldsymbol{u}(t), \quad \boldsymbol{x}_0 = \boldsymbol{x}(t_0) & (27) \\ \boldsymbol{y}(t) &= \boldsymbol{C}\boldsymbol{x}(t). & (28) \end{aligned}$$

The solution to the state equation is

$$\boldsymbol{x}(t) = e^{\boldsymbol{A}(t-t_0)}\boldsymbol{x}(t_0) + \int_{t_0}^{t} e^{\boldsymbol{A}(t-\tau)}\boldsymbol{B}\boldsymbol{u}(\tau)d\tau. \tag{29}$$

We assume that the input to the system is generated by a sample and zero-order device and has the form,

$$\boldsymbol{u}(t) = \boldsymbol{u}(k), \quad kh \leq t < (k+1)h.$$

Let $t_0 = kh$ and $t = (k+1)h$ and let us use shorthand notation,

$$\boldsymbol{x}[k] = \boldsymbol{x}(kh)$$

Then taking into account that $\boldsymbol{u}(t)$ is constant on the interval $[kh, (k+1)h)$ we represent (29) as

$$\begin{aligned}
\boldsymbol{x}[k+1] &= e^{\boldsymbol{A}h}\boldsymbol{x}[k] + \int_{kh}^{(k+1)h} e^{\boldsymbol{A}(kh+h-\tau)}\boldsymbol{B}\boldsymbol{u}[k]d\tau \\
&= e^{\boldsymbol{A}h}\boldsymbol{x}[k] + \int_{kh}^{(k+1)h} e^{\boldsymbol{A}(kh+h-\tau)}\boldsymbol{B}d\tau\, \boldsymbol{u}[k]. \qquad (30)
\end{aligned}$$

Consider now the second term on the right-hand side of the above equation. Let

$$\eta = kh + h - \tau.$$

Then we can represent (30) as

$$\begin{aligned}
\boldsymbol{x}[k+1] &= e^{\boldsymbol{A}h}\boldsymbol{x}[k] + \int_{kh}^{(k+1)h} e^{\boldsymbol{A}(kh+h-\tau)}\boldsymbol{B}d\tau\, \boldsymbol{u}[k] \\
&= e^{\boldsymbol{A}h}\boldsymbol{x}[k] + \int_{0}^{h} e^{\boldsymbol{A}\eta}\boldsymbol{B}d\eta\, \boldsymbol{u}[k] \\
&= \boldsymbol{A}_d\boldsymbol{x}[k] + \boldsymbol{B}_d\boldsymbol{u}[k], \qquad (31)
\end{aligned}$$

where

$$\boldsymbol{A}_d = e^{\boldsymbol{A}h} \quad \text{and} \quad \boldsymbol{B}_d = \int_0^h e^{\boldsymbol{A}\eta}\boldsymbol{B}d\eta.$$

The discrete output equation has the form

$$\boldsymbol{y}[k] = \boldsymbol{C}\boldsymbol{x}[k].$$

We point out that the above discrete model, employing the zero-order hold, is exact in the sense that the state, input, and output of this model are exactly the same at the sampling times as the state, input, and output of the underlying continuous-time model.

Example 10 We construct an exact linear discrete state-space model for the continuous system model,

$$\begin{aligned}
\dot{\boldsymbol{x}}(t) &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \boldsymbol{x}(k) + \begin{bmatrix} 2 \\ 1 \end{bmatrix} u(t) \\
y(t) &= \begin{bmatrix} 1 & 0 \end{bmatrix} \boldsymbol{x}(t).
\end{aligned}$$

The sampling period is $h = 1$ sec.

We first evaluate $\boldsymbol{A}_d = e^{\boldsymbol{A}h}$. Note that $\boldsymbol{A}^2 = \boldsymbol{O}$. Hence

$$\begin{aligned}\boldsymbol{A}_d &= \boldsymbol{I}_2 + \boldsymbol{A}h \\ &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.\end{aligned}$$

Next, we evaluate

$$\boldsymbol{B}_d = \int_0^h e^{\boldsymbol{A}\tau}\boldsymbol{B}d\tau.$$

We obtain

$$\begin{aligned}\boldsymbol{B}_d &= \int_0^h e^{\boldsymbol{A}\tau}\boldsymbol{B}d\tau \\ &= \int_0^h \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} d\tau \\ &= \int_0^1 \begin{bmatrix} 2+\tau \\ 1 \end{bmatrix} d\tau \\ &= \begin{bmatrix} 2\tau + \frac{1}{2}\tau^2 \\ \tau \end{bmatrix}\Bigg|_0^1 \\ &= \begin{bmatrix} 2.5 \\ 1 \end{bmatrix}.\end{aligned}$$

The discrete model then has the form

$$\begin{aligned}\boldsymbol{x}[k+1] &= \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\boldsymbol{x}[k] + \begin{bmatrix} 2.5 \\ 1 \end{bmatrix} \\ y[k] &= \begin{bmatrix} 1 & 0 \end{bmatrix}\boldsymbol{x}[k].\end{aligned}$$

6 Summary and Notes

7 Exercises

Exercise 1 Write a MATLAB program that animates the motion of the double pendulum system discussed in class. Obtain a plot of the motion of the pendulum tip versus time on the time interval [0, 5 sec]. Use the initial condition,

$$\begin{bmatrix} \theta_1(0) & \dot{\theta}_1(0) & \theta_2(0) & \dot{\theta}_2(0) \end{bmatrix} = \begin{bmatrix} 135^\circ & 0 & 180^\circ & 0 \end{bmatrix}.$$

Recommended step-size for solving differential equations is $dt = 0.0025$.

Exercise 2 Suppose that $\boldsymbol{x} = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}^\top$ is the representation of the given vector in the natural basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Find the representation of this vector in the basis $\{\boldsymbol{e}_1', \boldsymbol{e}_2', \boldsymbol{e}_3'\}$, where

$$\begin{aligned} \boldsymbol{e}_1' &= \boldsymbol{e}_1 + \boldsymbol{e}_3 \\ \boldsymbol{e}_2' &= \boldsymbol{e}_2 - \boldsymbol{e}_3 \\ \boldsymbol{e}_3' &= \boldsymbol{e}_3. \end{aligned}$$

Exercise 3 Prove Theorem 1.

Exercise 4 Find a similarity transformation that transforms

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

into its Jordan form. Do not fix λ; construct a transformation that works for any λ. Verify your calculations using MATLAB's Symbolic Toolbox.

Exercise 5 Calculate the solution for the system,

$$\dot{\boldsymbol{x}} = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u, \quad \boldsymbol{x}(0) = \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix},$$

where $u(t)$ is the unit step, that is, $u(t) = 1(t)$. Plot the system trajectories on the time interval $[0,\ 2]$ for $\boldsymbol{x}(0) = \boldsymbol{0}$.

Exercise 6 Construct an exact linear discrete state-space model for the system,

$$\begin{aligned} \dot{\boldsymbol{x}} &= \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u \\ y &= \begin{bmatrix} 0 & 1 \end{bmatrix} \boldsymbol{x} \end{aligned}$$

The sampling period is $h = 0.2$ sec. Verify you calculations using MATLAB's function `c2d`. Find the transfer function of the discrete systems and its step response. Plot the step response of the discrete system on the time interval [0, 2 sec].

Exercise 7 Use the Kronecker product to solve the following Sylvester's equation,

$$\boldsymbol{AX} + \boldsymbol{XB} = \boldsymbol{C},$$

where

$$\boldsymbol{A} = \begin{bmatrix} 2.7694 & 0.7254 & -0.2050 \\ -1.3499 & -0.0631 & -0.1241 \\ 3.0349 & 0.7147 & 1.4897 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 1.4090 & 0.7172 & 0.7269 & 0.8884 \\ 1.4172 & 1.6302 & -0.3034 & -1.1471 \\ 0.6715 & 0.4889 & 0.2939 & -1.0689 \\ -1.2075 & 1.0347 & -0.7873 & -0.8095 \end{bmatrix},$$

and

$$\boldsymbol{C} = \begin{bmatrix} -2.9443 & -0.7549 & -0.1022 & 0.3129 \\ 1.4384 & 1.3703 & -0.2414 & -0.8649 \\ 0.3252 & -1.7115 & 0.3192 & -0.0301 \end{bmatrix}.$$

Verify your solution using MATLAB's function `lyap`.

Exercise 8 **(i)** Does the rank of the matrix

$$\boldsymbol{Q} = \begin{bmatrix} \alpha^4 & \alpha^3 & \alpha^2 \\ \alpha^3 & \alpha^2 & \alpha \\ \alpha^2 & \alpha & 1 \end{bmatrix}, \quad \alpha \in \mathbb{R},$$

depend on the value of α? Is $\boldsymbol{Q}$ positive definite, positive semidefinite, negative definite, negative semidefinite, or indefinite? Justify your answer.

(ii) Find the range of the parameter γ for which the following quadratic form is positive definite;

$$f(x_1, x_2) = \gamma x_1^2 + x_2^2 + 4x_1x_2.$$

(iii) Find the range of values of the parameter ξ for which the function

$$f(x_1, x_2, x_3) = 2x_1x_3 - x_1^2 - x_2^2 - 5x_3^2 - 2\xi x_1x_2 - 4x_2x_3,$$

is negative semi-definite.

References

[1] C.-T. Chen. *Linear System Theory and Design.* Oxford University Press, New York, third edition, 1999.

[2] C. Moler and C. Van Loan. Nineteen dubious ways to compute the exponential of a matrix. *SIAM Review*, 20(4):801–836, October 1978.

[3] D. H. Owens. *Multivariable and Optimal Systems.* Academic Press, London, 1981.