6

Controllability and Observability: Special Forms

6.1 Introduction

In this chapter, important special forms for the state-space description of time-invariant systems are presented. These forms are obtained by means of similarity transformations and are designed to reveal those features of a system that are related to the properties of controllability and observability. In Section 6.2, special state-space forms that separate the controllable (observable) from the uncontrollable (unobservable) part of a system are presented. These forms, referred to as the standard forms for uncontrollable and unobservable systems, are very useful in establishing a number of results. In particular, these forms are used in Section 6.3 to derive alternative tests for controllability and observability and in Section 7.2 to relate state-space and input–output descriptions. In Section 6.4 the controller and observer state-space forms are introduced. These are useful in the study of state-space realizations in Chapter 8 and state feedback and state estimators in Chapter 9.

6.2 Standard Forms for Uncontrollable and Unobservable Systems

We consider time-invariant systems described by equations of the form

$$\dot{x} = Ax + Bu, \quad y = Cx + Du, \tag{6.1}$$

where $A \in R^{n\times n}$, $B \in R^{n\times m}$, $C \in R^{p\times n}$, and $D \in R^{p\times m}$. It was shown in the previous chapter that this system is state reachable if and only if the $n \times mn$ controllability matrix

$$\mathcal{C} \triangleq [B, AB, \ldots, A^{n-1}B] \tag{6.2}$$

has full row rank n; i.e., $\operatorname{rank} \mathcal{C} = n$. If the system is reachable (or controllable-from-the-origin), then it is also controllable (or controllable-to-the-origin), and vice versa (see Section 5.3.1).

It was also shown earlier that system (6.1) is state observable if and only if the $pn \times n$ observability matrix

$$\mathcal{O} \triangleq \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \tag{6.3}$$

has full column rank; i.e., rank $\mathcal{O} = n$. If the system is observable, then it is also constructible, and vice versa (see Section 5.4.1).

Similar results were also derived for discrete-time time-invariant systems described by equations of the form

$$x(k+1) = Ax(k) + Bu(k), \quad y(k) = Cx(k) + Du(k). \tag{6.4}$$

Again, rank $\mathcal{C} = n$ and rank $\mathcal{O} = n$ are the necessary and sufficient conditions for state reachability and observability, respectively. Reachability always implies controllability and observability always implies constructibility, as in the continuous-time case. However, in the discrete-time case, controllability does not necessarily imply reachability and constructibility does not imply observability, unless A is nonsingular (see Sections 5.3.2 and 5.4.2).

Next, we will introduce standard forms for unreachable and unobservable systems both for the continuous-time and the discrete-time time-invariant cases. These forms will be referred to as standard forms for *uncontrollable systems*, rather than unreachable systems, and standard forms for *unobservable systems*, respectively.

6.2.1 Standard Form for Uncontrollable Systems

If the system (6.1) [or (6.4)] is not completely reachable or controllable-from-the-origin, then it is possible to "separate" the controllable part of the system by means of an appropriate similarity transformation. This amounts to changing the basis of the state space so that all the vectors in the reachable subspace R_r have a certain structure. In particular, let rank $\mathcal{C} = n_r < n$; i.e., the pair (A, B) is not controllable. This implies that the subspace $R_r = \mathcal{R}(\mathcal{C})$ has dimension n_r. Let $\{v_1, v_2, \ldots, v_{n_r}\}$ be a basis for R_r. These n_r vectors can be, for example, any n_r linearly independent columns of $\mathcal{C}$. Define the $n \times n$ similarity transformation matrix

$$Q \triangleq [v_1, v_2, \ldots, v_{n_r}, Q_{n-n_r}], \tag{6.5}$$

where the $n \times (n - n_r)$ matrix Q_{n-n_r} contains $n - n_r$ linearly independent vectors chosen so that Q is nonsingular. There are many such choices. We are now in a position to prove the following result.

Lemma 6.1. *For (A, B) uncontrollable, there exists a nonsingular matrix Q such that*

$$\hat{A} = Q^{-1}AQ = \begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix} \quad \text{and} \quad \hat{B} = Q^{-1}B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \tag{6.6}$$

where $A_1 \in R^{n_r \times n_r}, B_1 \in R^{n_r \times m}$, and the pair (A_1, B_1) is controllable. The pair $(\hat{A}, \hat{B})$ is in the standard form for uncontrollable systems.

Proof. We need to show that

$$AQ = A[v_1, \ldots, v_{n_r}, Q_{n-n_r}] = [v_1, \ldots, v_{n_r}, Q_{n-n_r}] \begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix} = Q\hat{A}.$$

Since the subspace R_r is A-invariant (see Lemma 5.19), $Av_i \in R_r$, which can be written as a linear combination of only the n_r vectors in a basis of R_r. Thus, A_1 in $\hat{A}$ is an $n_r \times n_r$ matrix, and the $(n - n_r) \times n_r$ matrix below it in $\hat{A}$ is a zero matrix. Similarly, we also need to show that

$$B = [v_1, \ldots, v_{n_r}, Q_{n-n_r}] \begin{bmatrix} B_1 \\ 0 \end{bmatrix} = Q\hat{B}.$$

But this is true for similar reasons: The columns of B are in the range of $\mathcal{C}$ or in R_r. ■

The $n \times nm$ controllability matrix $\hat{\mathcal{C}}$ of $(\hat{A}, \hat{B})$ is

$$\hat{\mathcal{C}} = [\hat{B}, \hat{A}\hat{B}, \ldots, \hat{A}^{n-1}\hat{B}] = \begin{bmatrix} B_1 & A_1B_1 & \cdots & A_1^{n-1}B_1 \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{6.7}$$

which clearly has $\operatorname{rank} \hat{\mathcal{C}} = \operatorname{rank}[B_1, A_1B_1, \ldots, A_1^{n_r-1}B_1, \ldots, A_1^{n-1}B_1] = n_r$. Note that

$$\hat{\mathcal{C}} = Q^{-1}\mathcal{C}. \tag{6.8}$$

The range of $\hat{\mathcal{C}}$ is the controllable subspace of $(\hat{A}, \hat{B})$. It contains vectors only of the form $[\alpha^T, 0]^T$, where $\alpha \in R^{n_r}$. Since $\dim \mathcal{R}(\hat{\mathcal{C}}) = \operatorname{rank} \hat{\mathcal{C}} = n_r$, every vector of the form $[\alpha^T, 0]^T$ is a controllable (state) vector. In other words, the similarity transformation has changed the basis of R^n in such a manner so that all controllable vectors, expressed in terms of this new basis, have this very particular structure with zeros in the last $n - n_r$ entries.

Given system (6.1) [or (6.4)], if a new state $\hat{x}(t)$ is taken to be $\hat{x}(t) = Q^{-1}x(t)$, then

$$\dot{\hat{x}} = \hat{A}\hat{x} + \hat{B}u, \quad y = \hat{C}\hat{x} + \hat{D}u, \tag{6.9}$$

where $\hat{A} = Q^{-1}AQ$, $\hat{B} = Q^{-1}B$, $\hat{C} = CQ$, and $\hat{D} = D$ constitutes an equivalent representation (see Section 3.4.3). For Q as in Lemma 6.1, we obtain

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{bmatrix} = \begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u, y = [C_1, C_2] \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + Du, \tag{6.10}$$

where $\hat{x} = [\hat{x}_1^T, \hat{x}_2^T]^T$ with $\hat{x}_1 \in R^{n_r}$ and where (A_1, B_1) is controllable. The matrix $\widehat{C} = [C_1, C_2]$ does not have any particular structure. This representation is called a *standard form for the uncontrollable system.* The state equation can now be written as

$$\dot{\hat{x}}_1 = A_1\hat{x}_1 + B_1u + A_{12}\hat{x}_2, \dot{\hat{x}}_2 = A_2\hat{x}_2, \tag{6.11}$$

which shows that the input u does not affect the trajectory component $\hat{x}_2(t)$ at all, and therefore, $\hat{x}_2(t)$ is determined only by the value of its initial vector. The input u certainly affects $\hat{x}_1(t)$. Note also that the trajectory component $\hat{x}_1(t)$ is also influenced by $\hat{x}_2(t)$. In fact,

$$\hat{x}_1(t) = e^{A_1t}\hat{x}_1(0) + \int_0^t e^{A_1(t-\tau)}B_1u(\tau)d\tau + \left[\int_0^t e^{A_1(t-\tau)}A_{12}e^{A_2\tau}d\tau\right]\hat{x}_2(0). \tag{6.12}$$

The n_r eigenvalues of A_1 and the corresponding modes are the *controllable eigenvalues* and *controllable modes* of the pair (A, B) or of system (6.1) [or of (6.4)]. The $n - n_r$ eigenvalues of A_2 and the corresponding modes are the *uncontrollable eigenvalues* and *uncontrollable modes*, respectively.

It is interesting to observe that in the zero-state response of the system (zero initial conditions), the uncontrollable modes are completely absent. In particular, in the solution $x(t) = e^{At}x(0) + \int_0^t e^{A(t-\tau)}Bu(\tau)d\tau$ of $\dot{x} = Ax + Bu$, given $x(0)$, notice that

$$e^{A(t-\tau)}B = [Qe^{\widehat{A}(t-\tau)}Q^{-1}][Q\widehat{B}] = Q\begin{bmatrix} e^{A_1(t-\tau)}B_1 \\ 0 \end{bmatrix},$$

where A_1 [from (6.6)] contains only the controllable eigenvalues. Therefore, the input $u(t)$ cannot directly influence the uncontrollable modes. Note, however, that the uncontrollable modes do appear in the zero-input response $e^{At}x(0)$. The same observations can be made for discrete-time systems (6.4) where the quantity A^kB is of interest.

Example 6.2. Given $A = \begin{bmatrix} 0 & -1 & 1 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$, we wish to reduce system (6.1) to the standard form (6.6). Here

$$\mathcal{C} = [B, AB, A^2B] = \left[\begin{array}{cc|cc|cc} 1 & 0 & 0 & 1 & 0 & -1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & -1 & 0 & 1 \end{array}\right]$$

and $\operatorname{rank}\mathcal{C} = n_r = 2 < 3 = n$. Thus, the subspace $R_r = \mathcal{R}(\mathcal{C})$ has dimension $n_r = 2$, and a basis $\{v_1, v_2\}$ can be found by taking two linearly independent columns of $\mathcal{C}$, say, the first two, to obtain

$$Q = [v_1, v_2, Q_1] = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{array}\right].$$

The third column of Q was selected so that Q is nonsingular. Note that the first two columns of Q could have been the first and fourth columns of $\mathcal{C}$ instead, or any other two linearly independent vectors obtained as a linear combination of the columns in $\mathcal{C}$. For the above choice for Q, we have

$$\begin{aligned} \widehat{A} = Q^{-1}AQ &= \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 & 1 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & -1 & 1 \\ 1 & -1 & 0 \\ -2 & 4 & -2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \\ &= \left[\begin{array}{cc|c} 0 & 1 & 1 \\ 0 & -1 & 0 \\ \hline 0 & 0 & -2 \end{array}\right] = \left[\begin{array}{c|c} A_1 & A_{12} \\ \hline 0 & A_2 \end{array}\right], \\ \widehat{B} = Q^{-1}B &= \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ \hline 0 & 0 \end{array}\right] = \left[\begin{array}{c} B_1 \\ \hline 0 \end{array}\right], \end{aligned}$$

where (A_1, B_1) is controllable. The matrix A has three eigenvalues at $0, -1$, and -2. It is clear from $(\widehat{A}, \widehat{B})$ that the eigenvalues $0, -1$ are controllable (in A_1), whereas -2 is an uncontrollable eigenvalue (in A_2).

6.2.2 Standard Form for Unobservable Systems

The standard form for an unobservable system can be derived in a similar way as the standard form of uncontrollable systems. If the system (6.1) [or (6.4)] is not completely state observable, then it is possible to "separate" the unobservable part of the system by means of a similarity transformation. This amounts to changing the basis of the state space so that all the vectors in the unobservable subspace $R_{\bar{o}}$ have a certain structure.

As in the preceding discussion concerning systems or pairs (A, B) that are not completely controllable, we shall select a similarity transformation Q to reduce a pair (A, C), which is not completely observable, to a particular form. This can be accomplished in two ways. The simplest way is to invoke duality and to work with the pair $(A_D = A^T, B_D = C^T)$, which is not controllable (refer to the discussion of dual systems in Section 5.2.3). If Lemma 6.1 is applied, then

$$\widehat{A}_D = Q_D^{-1} A_D Q_D = \begin{bmatrix} A_{D1} & A_{D12} \\ 0 & A_{D2} \end{bmatrix}, \quad \widehat{B}_D = Q_D^{-1} B_D = \begin{bmatrix} B_{D1} \\ 0 \end{bmatrix},$$

where (A_{D1}, B_{D1}) is controllable.

Taking the dual again, we obtain the pair $(\widehat{A}, \widehat{C})$, which has the desired properties. In particular,

$$\begin{aligned} \widehat{A} &= \widehat{A}_D^T = Q_D^T A_D^T (Q_D^T)^{-1} = Q_D^T A (Q_D^T)^{-1} = \begin{bmatrix} A_{D1}^T & 0 \\ A_{D12}^T & A_{D2}^T \end{bmatrix}, \\ \widehat{C} &= \widehat{B}_D^T = B_D^T (Q_D^T)^{-1} = C (Q_D^T)^{-1} = [B_{D1}^T, 0], \end{aligned} \tag{6.13}$$

where (A_{D1}^T, B_{D1}^T) is completely observable by duality (see Lemma 5.7).

Example 6.3. Given $A = \begin{bmatrix} 0 & 1 & 0 \\ -1 & -2 & 1 \\ 1 & 1 & -1 \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix}$, we wish to reduce system (6.1) to the standard form (6.13). To accomplish this, let $A_D = A^T$ and $B_D = C^T$. Notice that the pair (A_D, B_D) is precisely the pair (A, B) of Example 6.2.

A pair (A, C) can of course also be reduced directly to the standard form for unobservable systems. This is accomplished in the following.

Consider the system (6.1) [or (6.4)] and the observability matrix $\mathcal{O}$ in (6.3). Let $\operatorname{rank} \mathcal{O} = n_o < n$; i.e., the pair (A, C) is not completely observable. This implies that the unobservable subspace $R_{\bar{o}} = \mathcal{N}(\mathcal{O})$ has dimension $n - n_o$. Let $\{v_1, \ldots, v_{n-n_o}\}$ be a basis for $R_{\bar{o}}$, and define an $n \times n$ similarity transformation matrix Q as

$$Q \triangleq [Q_{n_o}, v_1, \ldots, v_{n-n_o}], \tag{6.14}$$

where the $n \times n_o$ matrix Q_{n_o} contains n_o linearly independent vectors chosen so that Q is nonsingular. Clearly, there are many such choices.

Lemma 6.4. *For (A, C) unobservable, there is a nonsingular matrix Q such that*

$$\widehat{A} = Q^{-1} A Q = \begin{bmatrix} A_1 & 0 \\ A_{21} & A_2 \end{bmatrix} \quad \text{and} \quad \widehat{C} = CQ = [C_1,\ 0], \tag{6.15}$$

where $A_1 \in R^{n_o \times n_o}, C_1 \in R^{p \times n_o}$, and the pair (A_1, C_1) is observable. The pair $(\widehat{A}, \widehat{C})$ is in the standard form for unobservable systems.

Proof. We need to show that

$$AQ = A[Q_{n_0}, v_1, \ldots, v_{n-n_o}] = [Q_{n_o}, v_1, \ldots, v_{n-n_o}] \begin{bmatrix} A_1 & 0 \\ A_{21} & A_2 \end{bmatrix} = Q\widehat{A}.$$

Since the unobservable subspace $R_{\bar{o}}$ is A-invariant (see Lemma 5.49), $Av_i \in R_{\bar{o}}$, which can be written as a linear combination of only the $n - n_o$ vectors in a basis of $R_{\bar{o}}$. Thus, A_2 in $\widehat{A}$ is an $(n - n_o) \times (n - n_o)$ matrix, and the $n_o \times (n - n_o)$ matrix above it in $\widehat{A}$ is a zero matrix. Similarly, we also need to show that

$$CQ = C[Q_{n_o}, v_1, \ldots, v_{n-n_o}] = [C_1, 0] = \widehat{C}.$$

This is true since $Cv_i = 0$. ■

The $pn \times n$ observability matrix $\widehat{\mathcal{O}}$ of $(\widehat{A}, \widehat{C})$ is

$$\widehat{\mathcal{O}} = \begin{bmatrix} \widehat{C} \\ \widehat{C}\widehat{A} \\ \vdots \\ \widehat{C}\widehat{A}^{n-1} \end{bmatrix} = \begin{bmatrix} C_1 & 0 \\ C_1 A_1 & 0 \\ \vdots & \vdots \\ C_1 A_1^{n-1} & 0 \end{bmatrix}, \tag{6.16}$$

which clearly has

$$\operatorname{rank} \widehat{\mathcal{O}} = \operatorname{rank} \begin{bmatrix} C_1 \\ C_1 A_1 \\ \vdots \\ C_1 A_1^{n_o - 1} \\ \vdots \\ C_1 A_1^{n-1} \end{bmatrix} = n_o.$$

Note that

$$\widehat{\mathcal{O}} = \mathcal{O} Q. \tag{6.17}$$

The null space of $\widehat{\mathcal{O}}$ is the unobservable subspace of $(\widehat{A}, \widehat{C})$. It contains vectors only of the form $[0, \alpha^T]^T$, where $\alpha \in R^{n-n_0}$. Since $\dim \mathcal{N}(\widehat{\mathcal{O}}) = n - \operatorname{rank} \widehat{\mathcal{O}} = n - n_0$, every vector of the form $[0, \alpha^T]^T$ is an unobservable (state) vector. In other words, the similarity transformation has changed the basis of R^n in such a manner so that all unobservable vectors expressed in terms of this new basis have this very particular structure—zeros in the first n_o entries.

For Q chosen as in Lemma 6.4,

$$\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ A_{21} & A_2 \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u, y = [C_1, \ 0] \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} + Du, \tag{6.18}$$

where $\hat{x} = [\hat{x}_1^T, \hat{x}_2^T]^T$ with $\hat{x}_1 \in R^{n_o}$ and (A_1, C_1) is observable. The matrix $\widehat{B} = [B_1^T, B_2^T]^T$ does not have any particular form. This representation is called a *standard form for the unobservable system.*

The n_o eigenvalues of A_1 and the corresponding modes are called *observable eigenvalues* and *observable modes* of the pair (A, C) or of the system (6.1) [or of (6.4)]. The $n - n_o$ eigenvalues of A_2 and the corresponding modes are called *unobservable eigenvalues* and *unobservable modes*, respectively.

Notice that the trajectory component $\hat{x}(t)$, which is observed via the output y, is not influenced at all by $\hat{x}_2$, the trajectory of which is determined primarily by the eigenvalues of A_2.

The unobservable modes of the system are completely absent from the output. In particular, given $\dot{x} = Ax + Bu, y = Cx$ with initial state $x(0)$, we have

$$y(t) = Ce^{At}x(0) + \int_0^t Ce^{A(t-\tau)}Bu(\tau)d\tau$$

and $Ce^{At} = [\widehat{C}Q^{-1}][Qe^{\widehat{A}t}Q^{-1}] = [C_1 e^{A_1 t}, 0]Q^{-1}$, where A_1 [from (6.15)] contains only the observable eigenvalues. Therefore, the unobservable modes cannot be seen by observing the output. The same observations can be made for discrete-time systems where the quantity CA^k is of interest.

Example 6.5. Given $A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$ and $C = [1, 1]$, we wish to reduce system (6.1) to the standard form (6.15). To accomplish this, we compute $\mathcal{O} = \begin{bmatrix} C \\ CA \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -2 & -2 \end{bmatrix}$, which has rank $\mathcal{O} = n_o = 1 < 2 = n$. Therefore, the unobservable subspace $R_{\bar{o}} = \mathcal{N}(\mathcal{O})$ has dimension $n - n_o = 1$. In view of (6.14),

$$Q = [Q_1, v_1] = \left[\begin{array}{c|c} 0 & 1 \\ 1 & -1 \end{array}\right],$$

where $v_1 = [1, -1]^T$ is a basis for $R_{\bar{o}}$, and Q_1 was chosen so that Q is nonsingular. Then

$$\begin{aligned} \widehat{A} = Q^{-1}AQ &= \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} \\ &= \left[\begin{array}{c|c} -2 & 0 \\ \hline 1 & -1 \end{array}\right] = \begin{bmatrix} A_1 & 0 \\ A_{21} & A_2 \end{bmatrix}, \\ \widehat{C} = CQ &= [1,\ 1] \begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix} = [1,\ 0] = [C_1,\ 0], \end{aligned}$$

where (A_1, C_1) is observable. The matrix A has two eigenvalues at $-1, -2$. It is clear from $(\widehat{A}, \widehat{C})$ that the eigenvalue -2 is observable (in A_1), whereas -1 is an unobservable eigenvalue (in A_2).

6.2.3 Kalman's Decomposition Theorem

Lemmas 6.1 and 6.4 can be combined to obtain an equivalent representation of (6.1) where the reachable and observable parts of this system can readily be identified. We consider system (6.9) and proceed, in the following, to construct the $n \times n$ required similarity transformation matrix Q.

As before, we let n_r denote the dimension of the controllable subspace R_r; i.e., $n_r = \dim R_r = \dim \mathcal{R}(\mathcal{C}) = \operatorname{rank} \mathcal{C}$. The dimension of the unobservable subspace $R_{\bar{o}} = \mathcal{N}(\mathcal{O})$ is given by $n_{\bar{o}} = n - \operatorname{rank} \mathcal{O} = n - n_o$. Let $n_{r\bar{o}}$ be the dimension of the subspace $R_{r\bar{o}} \triangleq R_r \cap R_{\bar{o}}$, which contains all the state vectors $x \in R^n$ that are controllable but unobservable. We choose

$$Q \triangleq [v_1, \ldots, v_{n_r - n_{r\bar{o}}+1}, \ldots, v_{n_r}, Q_N, \hat{v}_1, \ldots, \hat{v}_{n_{\bar{o}} - n_{r\bar{o}}}], \tag{6.19}$$

where the n_r vectors in $\{v_1, \ldots, v_{n_r}\}$ form a basis for R_r. The last $n_{r\bar{o}}$ vectors $\{v_{n_r - n_{r\bar{o}}+1}, \ldots, v_{n_r}\}$ in the basis for R_r are chosen so that they form a basis for $R_{r\bar{o}} = R_r \cap R_{\bar{o}}$. The $n_{\bar{o}} - n_{r\bar{o}} = (n - n_o - n_{r\bar{o}})$ vectors $\{\hat{v}_1, \ldots, \hat{v}_{n_{\bar{o}}-n_{r\bar{o}}}\}$ are selected so that when taken together with the $n_{r\bar{o}}$ vectors $\{v_{n_r-n_{r\bar{o}}+1}, \ldots, v_{n_r}\}$ they form a basis for $R_{\bar{o}}$, the unobservable subspace. The remaining $N = n-(n_r+n_{\bar{o}}-n_{r\bar{o}})$ columns in Q_N are simply selected so that Q is nonsingular.

The following theorem is called the *Canonical Structure Theorem* or *Kalman's Decomposition Theorem.*

Theorem 6.6. *For (A, B) uncontrollable and (A, C) unobservable, there is a nonsingular matrix Q such that*

$$\widehat{A} = Q^{-1}AQ = \begin{bmatrix} A_{11} & 0 & A_{13} & 0 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ 0 & 0 & A_{33} & 0 \\ 0 & 0 & A_{43} & A_{44} \end{bmatrix}, \quad \widehat{B} = Q^{-1}B = \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{bmatrix}, \tag{6.20}$$
$$\widehat{C} = CQ = [C_1,\ 0,\ C_3,\ 0],$$

where

(i) (A_c, B_c) *with*

$$A_c \triangleq \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \quad \text{and} \quad B_c \triangleq \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

is controllable, where $A_c \in R^{n_r \times n_r}, B_c \in R^{n_r \times m}$*;*

(ii) (A_o, C_o) *with*

$$A_o \triangleq \begin{bmatrix} A_{11} & A_{13} \\ 0 & A_{33} \end{bmatrix} \quad \text{and} \quad C_o \triangleq [C_1, C_3]$$

is observable, where $A_o \in R^{n_o \times n_o}$ *and* $C_o \in R^{p \times n_o}$ *and where the dimensions of the matrices* $A_{ij}, B_i,$ *and* C_j *are as follows:*

$$\begin{aligned} &A_{11} : (n_r - n_{r\bar{o}}) \times (n_r - n_{r\bar{o}}), && A_{22} : n_{r\bar{o}} \times n_{r_{\bar{o}}}, \\ &A_{33} : (n - (n_r + n_{\bar{o}} - n_{r_{\bar{o}}})) \times && A_{44} : (n_{\bar{o}} - n_{r\bar{o}}) \times (n_{\bar{o}} - n_{r\bar{o}}), \\ &\qquad (n - (n_r + n_{\bar{o}} - n_{r_{\bar{o}}})), && \\ &B_1 : (n_r - n_{r\bar{o}}) \times m, && B_2 : n_{r\bar{o}} \times m, \\ &C_1 : p \times (n_r - n_{r\bar{o}}), && C_3 : p \times (n - (n_r + n_{\bar{o}} - n_{r\bar{o}})); \end{aligned}$$

(iii) the triple (A_{11}, B_1, C_1) *is such that* (A_{11}, B_1) *is controllable and* (A_{11}, C_1) *is observable.*

Proof. For details of the proof, refer to [6] and to [7], where further clarifications to [6] and an updated method of selecting Q are given. ■

The similarity transformation (6.19) has altered the basis of the state space in such a manner that the vectors in the controllable subspace R_r, the vectors in the unobservable subspace $R_{\bar{o}}$, and the vectors in the subspace $R_{r\bar{o}} \cap R_{\bar{o}}$ all have specific forms. To see this, we construct the controllability matrix $\widehat{\mathcal{C}} = [\widehat{B}, \ldots, \widehat{A}^{n-1}\widehat{B}]$ whose range is the controllable subspace and the observability matrix $\widehat{\mathcal{O}} = [\widehat{C}^T, \ldots, (\widehat{C}\widehat{A}^{n-1})^T]^T$, whose null space is the unobservable subspace. Then, all controllable states are of the form $[x_1^T, x_2^T, 0, 0]^T$, all the unobservable ones have the structure $[0, x_2^T, 0, x_4^T]^T$, and states of the form $[0, x_2^T, 0, 0]^T$ characterize $R_{r\bar{o}}$; i.e., they are controllable but unobservable.

Similarly to the previous two lemmas, the eigenvalues of $\widehat{A}$, or of A, are the eigenvalues of A_{11}, A_{22}, A_{33}, and A_{44}; i.e.,

$$|\lambda I - A| = |\lambda I - \widehat{A}| = |\lambda I - A_{11}||\lambda I - A_{22}||\lambda I - A_{33}||\lambda I - A_{44}|. \tag{6.21}$$

If we consider the representation $\{\widehat{A}, \widehat{B}, \widehat{C}, \widehat{D}\}$ given in (6.20), then

$$\begin{aligned}
\begin{bmatrix} \dot{\hat{x}}_1 \\ \dot{\hat{x}}_2 \\ \dot{\hat{x}}_3 \\ \dot{\hat{x}}_4 \end{bmatrix} &= \begin{bmatrix} A_{11} & 0 & A_{13} & 0 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ 0 & 0 & A_{33} & 0 \\ 0 & 0 & A_{43} & A_{44} \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{bmatrix} u, \\
y &= [C_1,\ 0,\ C_3,\ 0] \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{x}_3 \\ \hat{x}_4 \end{bmatrix} + Du.
\end{aligned} \tag{6.22}$$

This shows that the trajectory components corresponding to $\hat{x}_3$ and $\hat{x}_4$ are not affected by the input u. The modes associated with the eigenvalues of A_{33} and A_{44} determine the trajectory components for $\hat{x}_3$ and $\hat{x}_4$ (compare this with the results in Lemma 6.1). Similarly to Lemma 6.4, the trajectory components for $\hat{x}_2$ and $\hat{x}_4$ are not influenced by $\hat{x}_1$ and $\hat{x}_3$ (observed via y), and they are determined by the eigenvalues of A_{22} and A_{44}. The following is now apparent (see also Figure 6.1):

The eigenvalues of

A_{11} are controllable and observable,
A_{22} are controllable and unobservable,
A_{33} are uncontrollable and observable,
A_{44} are uncontrollable and unobservable.

Example 6.7. Given $A = \begin{bmatrix} 0 & -1 & 1 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$, and $C = [0,\ 1,\ 0]$, we wish to reduce system (6.1) to the canonical structure (or Kalman decomposition) form (6.20). The appropriate transformation matrix Q is given by (6.19). The matrix $\mathcal{C}$ was found in Example 6.2 and

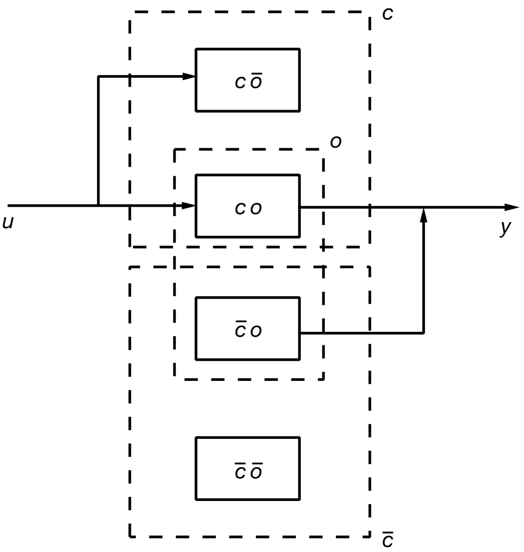

Figure 6.1. Canonical decomposition (c and $\bar{c}$ denote controllable and uncontrollable, respectively). The connections of the $c/\bar{c}$ and $o/\bar{o}$ parts of the system to the input and output are emphasized. Note that the impulse response (transfer function) of the system, which is an input–output description only, represents the part of the system that is both controllable and observable (see Chapter 7).

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -2 & 1 \\ -2 & 4 & -2 \end{bmatrix}.$$

A basis for $R_{\bar{o}} = \mathcal{N}(\mathcal{O})$ is $\{(1,\ 0,\ -1)^T\}$. Note that $n_r = 2, n_{\bar{o}} = 1$, and $n_{r\bar{o}} = 1$. Therefore,

$$Q = [v_1, v_2, Q_N] = \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{array}\right]$$

is an appropriate similarity matrix (check that $\det Q \neq 0$). We compute

$$\hat{A} = Q^{-1}AQ = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 & 1 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix}$$

$$= \left[\begin{array}{c|c|c} 0 & 0 & 1 \\ \hline 0 & -1 & 0 \\ \hline 0 & 0 & -2 \end{array}\right] = \begin{bmatrix} A_{11} & 0 & A_{13} \\ A_{21} & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{bmatrix},$$

$$\hat{B} = Q^{-1}B = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \left[\begin{array}{cc} 1 & 1 \\ \hline 0 & -1 \\ \hline 0 & 0 \end{array}\right] = \begin{bmatrix} B_1 \\ B_2 \\ 0 \end{bmatrix},$$

and

$$\widehat{C} = CQ = [0,\ 1,\ 0] \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix} = [1, 0, 0] = [C_1,\ 0,\ C_3].$$

The eigenvalue 0 (in A_{11}) is controllable and observable, the eigenvalue -1 (in A_{22}) is controllable and unobservable and the eigenvalue -2 (in A_{33}) is uncontrollable and observable. There are no eigenvalues that are both uncontrollable and unobservable.

6.3 Eigenvalue/Eigenvector Tests for Controllability and Observability

There are tests for controllability and observability for both continuous- and discrete-time time-invariant systems that involve the eigenvalues and eigenvectors of A. Some of these criteria are called PBH tests, after the initials of the codiscoverers (Popov–Belevitch–Hautus) of these tests. These tests are useful in theoretical analysis, and in addition, they are also attractive as computational tools.

Theorem 6.8. *(i) The pair (A, B) is uncontrollable if and only if there exists a $1 \times n$ (in general) complex vector $\hat{v}_i \neq 0$ such that*

$$\hat{v}_i[\lambda_i I - A, B] = 0, \tag{6.23}$$

where λ_i is some complex scalar.

(ii) The pair (A, C) is unobservable if and only if there exists an $n \times 1$ (in general) complex vector $v_i \neq 0$ such that

$$\begin{bmatrix} \lambda_i I - A \\ C \end{bmatrix} v_i = 0, \tag{6.24}$$

where λ_i is some complex scalar.

Proof. Only part (i) will be considered since (ii) can be proved using a similar argument or, directly, by duality arguments.

(*Sufficiency*) Assume that (6.23) is satisfied. In view of $\hat{v}_i A = \lambda_i \hat{v}_i$ and $\hat{v}_i B = 0, \hat{v}_i AB = \lambda_i \hat{v}_i B = 0$ and $\hat{v}_i A^k B = 0 \quad k = 0, 1, 2, \ldots$. Therefore, $\hat{v}_i \mathcal{C} = \hat{v}_i[B, AB, \ldots, A^{n-1}B] = 0$, which shows that (A, B) is not completely controllable.

(*Necessity*) Let (A, B) be uncontrollable and assume without loss of generality the standard form for A and B given in Lemma 6.1. We will show that there exist λ_i and $\hat{v}_i$ so that (6.23) holds. Let λ_i be an uncontrollable eigenvalue, and let $\hat{v}_i = [0, \alpha], \alpha^T \in C^{n-n_r}$, where $\alpha(\lambda_i I - A_2) = 0$; i.e., α is a left eigenvector of A_2 corresponding to λ_i. Then $\hat{v}_i[\lambda_i I - A, B] = [0, \alpha(\lambda_i I - A_2), 0] = 0$; i.e., (6.23) is satisfied. ■

Corollary 6.9. *(i)* λ_i *is an uncontrollable eigenvalue of* (A, B) *if and only if there exists a* $1 \times n$ *(in general) complex vector* $\hat{v}_i \neq 0$ *that satisfies (6.23).*
(ii) λ_i *is an unobservable eigenvalue of* (A, C) *if and only if there exists an* $n \times 1$ *(in general) complex vector* $v_i \neq 0$ *that satisfies (6.24).*

Proof. See [1, p. 273, Corollary 4.6]. ■

Example 6.10. Given are $A = \begin{bmatrix} 0 & -1 & 1 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$, and $C = [0,\ 1,\ 0]$, as in Example 6.7. The matrix A has three eigenvalues, $\lambda_1 = 0, \lambda_2 = -1$, and $\lambda_3 = -2$, with corresponding right eigenvectors $v_1 = [1,\ 1,\ 1]^T$, $v_2 = [1,\ 0,\ -1]^T$, $v_3 = [1,\ 1,\ -1]^T$ and with left eigenvectors $\hat{v}_1 = [1/2,\ 0,\ 1/2]$, $\hat{v}_2 = [1,\ -1,\ 0]$, and $\hat{v}_3 = [-1/2,\ 1,\ -1/2]$, respectively.

In view of Corollary 6.9, $\hat{v}_1 B = [1, 1] \neq 0$ implies that $\lambda_1 = 0$ is controllable. This is because $\hat{v}_1$ is the only nonzero vector (within a multiplication by a nonzero scalar) that satisfies $\hat{v}_1(\lambda_1 I - A) = 0$, and so $\hat{v}_1 B \neq 0$ implies that the only 1×3 vector α that satisfies $\alpha[\lambda_1 I - A, B] = 0$ is the zero vector, which in turn implies that λ_1 is controllable in view of (i) of Corollary 6.9. For similar reasons $Cv_1 = 1 \neq 0$ implies that $\lambda_1 = 0$ is observable; see (ii) of Corollary 6.9. Similarly, $\hat{v}_2 B = [0, -1] \neq 0$ implies that $\lambda_2 = -1$ is controllable, and $Cv_2 = 0$ implies that $\lambda_2 = -1$ is *unobservable*. Also, $\hat{v}_3 B = [0, 0]$ implies that $\lambda_3 = -2$ is *uncontrollable*, and $Cv_3 = 1 \neq 0$ implies that $\lambda_3 = -2$ is observable. These results agree with the results derived in Example 6.7.

Corollary 6.11. *(*Rank Tests*)*

(ia) The pair (A, B) *is controllable if and only if*

$$\text{rank}[\lambda I - A, B] = n \tag{6.25}$$

for all complex numbers λ*, or for all* n *eigenvalues* λ_i *of* A*.*

(ib) λ_i *is an uncontrollable eigenvalue of* A *if and only if*

$$\text{rank}[\lambda_i I - A, B] < n. \tag{6.26}$$

(iia) The pair (A, C) *is observable if and only if*

$$\text{rank} \begin{bmatrix} \lambda I - A \\ C \end{bmatrix} = n \tag{6.27}$$

for all complex numbers λ*, or for all* n *eigenvalues* λ_i*.*

(iib) λ_i *is an unobservable eigenvalue of* A *if and only if*

$$\text{rank} \begin{bmatrix} \lambda_i I - A \\ C \end{bmatrix} < n. \tag{6.28}$$

Proof. The proofs follow in a straightforward manner from Theorem 6.8. Notice that the only values of λ that can possibly reduce the rank of $[\lambda I - A, B]$ are the eigenvalues of A. ■

Example 6.12. If in Example 6.10 the eigenvalues $\lambda_1, \lambda_2, \lambda_3$ of A are known, but the corresponding eigenvectors are not, consider the system matrix

$$P(s) = \begin{bmatrix} sI - A & B \\ -C & 0 \end{bmatrix} = \left[\begin{array}{ccc|cc} s & 1 & -1 & 1 & 0 \\ -1 & s+2 & -1 & 1 & 1 \\ 0 & -1 & s+1 & 1 & 2 \\ \hline 0 & -1 & 0 & 0 & 0 \end{array}\right]$$

and determine $\text{rank}[\lambda_i I - A, B]$ and $\text{rank}\begin{bmatrix} \lambda_i I - A \\ C \end{bmatrix}$. Notice that

$$\text{rank}\begin{bmatrix} sI - A \\ C \end{bmatrix}_{s=\lambda_2} = \text{rank}\begin{bmatrix} -1 & 1 & -1 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} = 2 < 3 = n$$

and

$$\text{rank}[sI - A, B]_{s=\lambda_3} = \text{rank}\begin{bmatrix} -2 & 2 & -1 & 1 & 0 \\ -1 & 0 & -1 & 1 & 1 \\ 0 & -1 & -1 & 1 & 2 \end{bmatrix} = 2 < 3 = n.$$

In view of Corollary 6.11, $\lambda_2 = -1$ is unobservable and $\lambda_3 = -2$ is uncontrollable.

6.4 Controller and Observer Forms

It has been seen several times in this book that equivalent representations of systems

$$\dot{x} = Ax + Bu, \quad y = Cx + Du, \tag{6.29}$$

given by the equations

$$\dot{\hat{x}} = \widehat{A}\hat{x} + \widehat{B}u, \quad y = \widehat{C}\hat{x} + \widehat{D}u, \tag{6.30}$$

where $\hat{x} = Px, \widehat{A} = PAP^{-1}$, $\widehat{B} = PB$, $\widehat{C} = CP^{-1}$, and $\widehat{D} = D$ may offer advantages over the original representation when P (or $Q = P^{-1}$) is chosen in an appropriate manner. This is the case when P (or Q) is such that the new basis of the state space provides a natural setting for the properties of interest. This section shows how to select Q when (A, B) is controllable [or (A, C) is observable] to obtain the controller and observer forms. These special forms are very useful, in realizations discussed in Chapter 8 and especially when studying state-feedback control (and state observers) discussed in Chapter 9. They are also very useful in establishing a convenient way to transition between state-space representations and another very useful class of equivalent internal representations, the polynomial matrix representations.

Controller forms are considered first. Observer forms can of course be obtained directly in a similar manner to the controller forms, or they may be obtained by duality. This is addressed in the latter part of this section.

6.4.1 Controller Forms

The controller form is a particular system representation where both matrices (A, B) have a certain special structure. Since in this case A is in the companion form, the controller form is sometimes also referred to as the *controllable companion form*. Consider the system

$$\dot{x} = Ax + Bu, \quad y = Cx + Du, \tag{6.31}$$

where $A \in R^{n\times n}$, $B \in R^{n\times m}$, $C \in R^{p\times n}$, and $D \in R^{p\times m}$ and let (A, B) be controllable. Then $\operatorname{rank}\mathcal{C} = n$, where

$$\mathcal{C} = [B, AB, \ldots, A^{n-1}B]. \tag{6.32}$$

Assume that

$$\operatorname{rank} B = m \le n. \tag{6.33}$$

Under these assumptions, $\operatorname{rank}\mathcal{C} = n$ and $\operatorname{rank} B = m$. We will show how to obtain an equivalent pair $(\widehat{A}, \widehat{B})$ in controller form, first for the single-input case ($m = 1$) and then for the multi-input case ($m > 1$). Before this is accomplished, we discuss how to deal with two special cases that do not satisfy the above assumptions that $\operatorname{rank} B = m$ and that (A, B) is controllable.

1. If the m columns of B are not linearly independent ($\operatorname{rank} B = r < m$), then there exists an $m \times m$ nonsingular matrix K so that $BK = [B_r, 0]$, where the r columns of B_r are linearly independent ($\operatorname{rank} B_r = r$). Note that $\dot{x} = Ax + Bu = Ax + (BK)(K^{-1}u) = Ax + [B_r, 0]\begin{bmatrix} u_r \\ u_{m-r} \end{bmatrix} = Ax + B_r u_r$, which shows that when $\operatorname{rank} B = r < m$ the same input action to the system can be accomplished by only r inputs, instead of m inputs. The pair (A, B_r), which is controllable when (A, B) is controllable, can now be reduced to controller form, using the method described below.
2. When (A, B) is not completely controllable, then a two-step approach can be taken. First, the controllable part is isolated (see Subsection 6.2.1) and then is reduced to the controller form, using the methods of this section. In particular, consider the system $\dot{x} = Ax + Bu$ with $A \in R^{n\times n}, B \in R^{n\times m}$, and $\operatorname{rank} B = m$. Let $\operatorname{rank}[B, AB, \ldots, A^{n-1}B] = n_r < n$. Then

there exists a transformation P_1 such that $P_1AP_1^{-1} = \begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix}$ and $P_1B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}$, where $A_1 \in R^{n_r \times n_r}, B_1 \in R^{n_r \times m}$, and (A_1, B_1) is controllable (Subsection 6.2.1). Since (A_1, B_1) is controllable, there exists a transformation P_2 such that $P_2A_1P_2^{-1} = A_{1c}$, and $P_2B_1 = B_{1c}$, where A_{1c}, B_{1c} is in controller form, defined below. Combining, we obtain

$$PAP^{-1} = \begin{bmatrix} A_{1c} & P_2A_{12} \\ 0 & A_2 \end{bmatrix}, \quad \text{and} \quad PB = \begin{bmatrix} B_{1c} \\ 0 \end{bmatrix} \tag{6.34}$$

[where $A_{1c} \in R^{n_r \times n_r}, B_{1c} \in R^{n_r \times m}$, and (A_{1c}, B_{1c}) is controllable], which is in controller form. Note that

$$P = \begin{bmatrix} P_2 & 0 \\ 0 & I \end{bmatrix} P_1. \tag{6.35}$$

Single-Input Case ($m = 1$)

The representation $\{A_c, B_c, C_c, D_c\}$ in controller form is given by $A_c \triangleq \widehat{A} = PAP^{-1}$ and $B_c \triangleq \widehat{B} = PB$ with

$$A_c = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ -\alpha_0 & -\alpha_1 & \cdots & -\alpha_{n-1} \end{bmatrix}, \quad B_c = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \tag{6.36}$$

where the coefficients α_i are the coefficients of the characteristic polynomial $\alpha(s)$ of A; that is,

$$\alpha(s) \triangleq \det(sI - A) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s + \alpha_0. \tag{6.37}$$

Note that $C_c \triangleq \widehat{C} = CP^{-1}$ and $D_c = D$ do not have any particular structure. The structure of (A_c, B_c) is very useful (in control problems), and the representation $\{A_c, B_c, C_c, D_c\}$ shall be referred to as the *controller form* of the system. The similarity transformation matrix P is obtained as follows. The controllability matrix $\mathcal{C} = [B, AB, \ldots, A^{n-1}B]$ is in this case an $n \times n$ nonsingular matrix. Let $\mathcal{C}^{-1} = \begin{bmatrix} \times \\ q \end{bmatrix}$, where q is the nth row of $\mathcal{C}^{-1}$ and $\times$ indicates the remaining entries of $\mathcal{C}^{-1}$. Then

$$P \triangleq \begin{bmatrix} q \\ qA \\ \cdots \\ qA^{n-1} \end{bmatrix}. \tag{6.38}$$

To show that $PAP^{-1} = A_c$ and $PB = B_c$ given in (6.36), note first that $qA^{i-1}B = 0$ $i = 1, \dots, n-1$ and $qA^{n-1}\,B = 1$. This can be verified from the definition of q, which implies that $q\,\mathcal{C} = [0, 0, \dots, 1]$. Now

$$P\mathcal{C} = P[B, AB, \dots, A^{n-1}B] = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 1 \\ 0 & 0 & \cdots & 1 & \times \\ \vdots & 1 & & \vdots & \vdots \\ 1 & \times & \cdots & \times & \times \end{bmatrix} = \mathcal{C}_c, \tag{6.39}$$

which implies that $|P\mathcal{C}| = |P|\,|\mathcal{C}| \neq 0$ or that $|P| \neq 0$. Therefore, P qualifies as a similarity transformation matrix. In view of (6.39), $PB = [0, 0, \dots, 1]^T = B_c$. Furthermore,

$$A_cP = \begin{bmatrix} qA \\ \vdots \\ qA^{n-1} \\ qA^n \end{bmatrix} = PA, \tag{6.40}$$

where in the last row of A_cP, the relation $-\sum_{i=0}^{n-1} \alpha_i A^i = A^n$ was used [which is the Cayley–Hamilton Theorem, namely, $\alpha(A) = 0$].

Example 6.13. Let $A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$. Since $n = 3$ and $|sI - A| = (s+1)(s-1)(s+2) = s^3 + 2s^2 - s - 2$, $\{A_c, B_c\}$ in controller form is given by

$$A_c = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 1 & -2 \end{bmatrix} \text{ and } B_c = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

The transformation matrix P that reduces (A, B) to $(A_c = PAP^{-1}, B_c = PB)$ is now derived. We have

$$\mathcal{C} = [B, AB, A^2B] = \begin{bmatrix} 1 & -1 & 1 \\ -1 & -1 & -1 \\ 1 & -2 & 4 \end{bmatrix} \quad \text{and} \quad \mathcal{C}^{-1} = \begin{bmatrix} 1 & -1/3 & -1/3 \\ -1/2 & -1/2 & 0 \\ -1/2 & -1/6 & 1/3 \end{bmatrix}.$$

The third (the nth) row of $\mathcal{C}^{-1}$ is $q = [-1/2, -1/6, 1/3]$, and therefore,

$$P \triangleq \begin{bmatrix} q \\ qA \\ qA^2 \end{bmatrix} = \begin{bmatrix} -1/2 & -1/6 & 1/3 \\ 1/2 & -1/6 & -2/3 \\ -1/2 & -1/6 & 4/3 \end{bmatrix}.$$

It can now easily be verified that $A_c = PAP^{-1}$, or

$$A_cP = \begin{bmatrix} 1/2 & -1/6 & -2/3 \\ -1/2 & -1/6 & -2/3 \\ 1/2 & -1/6 & -8/3 \end{bmatrix} = PA,$$

and that $B_c = PB$.

An alternative form to (6.36) is

$$A_{c1} = \begin{bmatrix} -\alpha_{n-1} & \cdots & -\alpha_1 & -\alpha_0 \\ 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}, \quad B_{c1} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{6.41}$$

which is obtained if the similarity transformation matrix is taken to be

$$P_1 \triangleq \begin{bmatrix} qA^{n-1} \\ \vdots \\ qA \\ q \end{bmatrix}, \tag{6.42}$$

i.e., by reversing the order of the rows of P in (6.38). (See Exercise 6.5 and Example 6.14.)

In the above, A_c is a companion matrix of the form $\begin{bmatrix} 0 & I \\ \times & \times \end{bmatrix}$ or $\begin{bmatrix} \times & \times \\ I & 0 \end{bmatrix}$. It could also be of the form $\begin{bmatrix} 0 & \times \\ I & \times \end{bmatrix}$ or $\begin{bmatrix} \times & 0 \\ \times & I \end{bmatrix}$ with coefficients $-[\alpha_0, \ldots, \alpha_{n-1}]^T$ in the last or the first column. It is shown here, for completeness, how to determine controller forms where A_c are such companion matrices. In particular, if

$$Q_2 = P_2^{-1} = [B, AB, \ldots, A^{n-1}B] = \mathcal{C}, \tag{6.43}$$

then

$$A_{c2} = Q_2^{-1}AQ_2 = \begin{bmatrix} 0 & \cdots & 0 & -\alpha_0 \\ 1 & \cdots & 0 & -\alpha_1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & -\alpha_{n-1} \end{bmatrix}, B_{c2} = Q_2^{-1}B = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{6.44}$$

Also, if

$$Q_3 = P_3^{-1} = [A^{n-1}B, \ldots, B], \tag{6.45}$$

then

$$A_{c3} = Q_3^{-1}AQ_3 = \begin{bmatrix} -\alpha_{n-1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\alpha_1 & 0 & \cdots & 1 \\ -\alpha_0 & 0 & \cdots & 0 \end{bmatrix}, B_{c3} = Q_3^{-1}B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{6.46}$$

(A_c, B_c) in (6.44) and (6.46) are also in controller canonical or controllable companion form. (See also Exercise 6.5 and Example 6.14.)

Example 6.14. Let $A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$, as in Example 6.13. Alternative controller forms can be derived for different P. In particular, if

(i) $P = P_1 = \begin{bmatrix} qA^2 \\ qA \\ q \end{bmatrix} = \begin{bmatrix} -1/2 & -1/6 & 4/3 \\ 1/2 & -1/6 & -2/3 \\ -1/2 & -1/6 & 1/3 \end{bmatrix}$, as in (6.42) ($\mathcal{C}, \mathcal{C}^{-1}$, and q were found in Example 6.13), then

$$A_{c1} = \begin{bmatrix} -2 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad B_{c1} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

as in (6.41). Note that in the present case $A_{c1}P_1 = \begin{bmatrix} 1/2 & -1/6 & -8/3 \\ -1/2 & -1/6 & 4/3 \\ 1/2 & -1/6 & -2/3 \end{bmatrix} = P_1 A, \quad B_{c1} = P_1 B.$

(ii) $Q_2 = \mathcal{C} = \begin{bmatrix} 1 & -1 & 1 \\ -1 & -1 & -1 \\ 1 & -2 & 4 \end{bmatrix}$, as in (6.43). Then

$$A_{c2} = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix}, \quad B_{c2} = Q_2^{-1} B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

as in (6.44).

(iii) $Q_3 = [A^2 B, AB, B] = \begin{bmatrix} 1 & -1 & 1 \\ -1 & -1 & -1 \\ 4 & -2 & 1 \end{bmatrix}$, as in (6.45). Then

$$A_{c3} = \begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \\ 2 & 0 & 0 \end{bmatrix}, \quad B_{c3} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

as in (6.46). Note that $Q_3 A_{c3} = \begin{bmatrix} -1 & 1 & -1 \\ -1 & -1 & -1 \\ -8 & 4 & -2 \end{bmatrix} = AQ_3, \quad Q_3 B_{c3} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = B.$

Multi-Input Case ($m > 1$)

In this case, the $n \times mn$ matrix $\mathcal{C}$ given in (6.32) is not square, and there are typically many sets of n columns of $\mathcal{C}$ that are linearly independent ($\operatorname{rank} \mathcal{C} = n$). Depending on which columns are chosen and in what order, different controller forms (controllable companion forms) are derived. Note that in the case when $m = 1$, four different controller forms were derived, even though there was only one set of n linearly independent columns. In the present case there are many more such choices. The form that will be used most often in the following is a generalization of (A_c, B_c) given in (6.36). Further discussion including derivation and alternative forms may be found in [1, Subsection 3.4D].

Let $\widehat{A} = PAP^{-1}$ and $\widehat{B} = PB$, where P is constructed as follows. Consider

$$\begin{aligned}\mathcal{C} &= [B, AB, \ldots, A^{n-1}B] \\ &= [b_1, \ldots, b_m, Ab_1, \ldots, Ab_m, \ldots, A^{n-1}b_1, \ldots, A^{n-1}b_m],\end{aligned} \tag{6.47}$$

where the $b_1, \ldots, b_m$ are the m columns of B. Select, starting from the left and moving to the right, the first n independent columns (rank $\mathcal{C} = n$). Reorder these columns by taking first b_1, Ab_1, A^2b_1, etc., until all columns involving b_1 have been taken; then take b_2, Ab_2, etc.; and lastly, take b_m, Ab_m, etc., to obtain

$$\bar{\mathcal{C}} \triangleq [b_1, Ab_1, \ldots, A^{\mu_1 - 1}b_1, \ldots, b_m, \ldots, A^{\mu_m - 1}b_m], \tag{6.48}$$

an $n \times n$ matrix. The integer μ_i denotes the number of columns involving b_i in the set of the first n linearly independent columns found in $\mathcal{C}$ when moving from left to right.

Definition 6.15. *The m integers μ_i, $i = 1, \ldots, m$, are the* controllability indices of the system, *and $\mu \triangleq \max \mu_i$ is called* the controllability index of the system. *Note that*

$$\sum_{i=1}^{m} \mu_i = n \quad \textit{and} \quad m\mu \geq n. \tag{6.49}$$

■

An alternative but equivalent definition for μ is that μ is the minimum integer k such that

$$\operatorname{rank}[B, AB, \ldots, A^{k-1}B] = n. \tag{6.50}$$

Notice that in (6.48) all columns of B are always present since rank $B = m$. This implies that $\mu_i \geq 1$ for all i. Notice further that if $A^k b_i$ is present, then $A^{k-1}b_i$ must also be present.

Now define

$$\sigma_k \triangleq \sum_{i=1}^{k} \mu_i, \quad k = 1, \ldots, m; \tag{6.51}$$

i.e., $\sigma_1 = \mu_1, \sigma_2 = \mu_1 + \mu_2, \ldots, \sigma_m = \mu_1 + \cdots + \mu_m = n$. Also, consider $\bar{\mathcal{C}}^{-1}$ and let q_k, where $q_k^T \in R^n$, $k = 1, \ldots, m$, denote its $\sigma_k{}^{th}$ row; i.e.,

$$\bar{\mathcal{C}}^{-1} = [\times, \ldots, \times, q_1^T \vdots \cdots \vdots \times, \ldots, \times, q_m^T]^T. \tag{6.52}$$

Next, define

$$P \triangleq \begin{bmatrix} q_1 \\ q_1 A \\ \vdots \\ q_1 A^{\mu_1 - 1} \\ \cdots \\ \vdots \\ \cdots \\ q_m \\ q_m A \\ \vdots \\ q_m A^{\mu_m - 1} \end{bmatrix}. \tag{6.53}$$

It can now be shown that $PAP^{-1} = A_c$ and $PB = B_c$ with

$$A_c = [A_{ij}], \qquad i, j = 1, \ldots, m,$$

$$A_{ii} = \begin{bmatrix} 0 & \\ \vdots & I_{\mu_i - 1} \\ 0 & \\ \times & \times \cdots \times \end{bmatrix} \in R^{\mu_i \times \mu_i},\ i = j, \quad A_{ij} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \\ \times & \times \cdots & \times \end{bmatrix} \in R^{\mu_i \times \mu_j},\ i \neq j,$$

and

$$B_c = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{bmatrix}, \qquad B_i = \begin{bmatrix} 0 \cdots 0 & 0 & \cdots & 0 \\ \vdots \quad \vdots & \vdots & & \vdots \\ 0 \cdots 0 & 1 & \times \cdots & \times \end{bmatrix} \in R^{\mu_i \times m}, \tag{6.54}$$

where the 1 in the last row of B_i occurs at the ith column location, $i = 1, \ldots, m$, and $\times$ denotes nonfixed entries. Note that $C_c = CP^{-1}$ does not have any particular structure. The expression (6.54) is a very useful form (in control problems) and shall be referred to as the *controller form* of the system. The derivation of this result is discussed in [1, Subsection 3.4D] .

Example 6.16. Given are $A \in R^{n \times n}$ and $B \in R^{n \times m}$ with (A, B) controllable and with $\operatorname{rank} B = m$. Let $n = 4$ and $m = 2$. Then there must be two controllability indices μ_1 and μ_2 such that $n = 4 = \sum_{i=1}^{2} \mu_i = \mu_1 + \mu_2$. Under these conditions, there are three possibilities:

(i) $\mu_1 = 2, \mu_2 = 2$,

$$A_c = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ \times & \times & \times & \times \\ \hline 0 & 0 & 0 & 1 \\ \times & \times & \times & \times \end{array}\right], \quad B_c = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \left[\begin{array}{cc} 0 & 0 \\ 1 & \times \\ \hline 0 & 0 \\ 0 & 1 \end{array}\right].$$

(ii) $\mu_1 = 1, \mu_2 = 3$,

$$A_c = \left[\begin{array}{cc|cc} \times & \times & \times & \times \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \times & \times & \times & \times \end{array}\right], \quad B_c = \left[\begin{array}{cc} 1 & \times \\ \hline 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{array}\right].$$

(iii) $\mu_1 = 3$, $\mu_2 = 1$,

$$A_c = \left[\begin{array}{ccc|c} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \times & \times & \times & \times \\ \hline \times & \times & \times & \times \end{array}\right], \quad B_c = \left[\begin{array}{cc} 0 & 0 \\ 0 & 0 \\ 1 & \times \\ \hline 0 & 1 \end{array}\right].$$

It is possible to write A_c, B_c in a systematic and perhaps more transparent way. In particular, notice that A_c, B_c in (6.54) can be expressed as

$$A_c = \bar{A}_c + \bar{B}_c A_m, \quad B_c = \bar{B}_c B_m, \tag{6.55}$$

where $\bar{A}_c = \text{block diag}[\bar{A}_{11}, \bar{A}_{22}, \ldots, \bar{A}_{mm}]$ with

$$\bar{A}_{ii} = \begin{bmatrix} 0 & \\ \vdots & I_{\mu_i - 1} \\ 0 & \\ 0 & 0 \cdots 0 \end{bmatrix} \in R^{\mu_i \times \mu_i}, \; \bar{B}_c = \text{block diag}\left(\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in R^{\mu_i \times 1}, \quad i = 1, \ldots, m \right),$$

and $A_m \in R^{m \times n}$ and $B_m \in R^{m \times m}$ are some appropriate matrices with $\sum_{i=1}^m \mu_i = n$. Note that the matrices $\bar{A}_c, \bar{B}_c$ are completely determined by the m controllability indices μ_i, $i = 1, \ldots, m$. The matrices A_m and B_m consist of the σ_1th, σ_2th, $\ldots, \sigma_m$th rows of A_c (entries denoted by $\times$) and the same rows of B_c, respectively [see (6.57) and (6.58) below].

Example 6.17. Let $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 2 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}$. To determine the controller form (6.54), consider

$$\mathcal{C} = [B, AB, A^2B] = [b_1, b_2, Ab_1, Ab_2, A^2b_1, A^2b_2] = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 2 & 2 \\ 0 & 0 & 2 & 2 & -2 & -2 \end{bmatrix},$$

where rank $\mathcal{C} = 3 = n$; i.e., (A, B) is controllable. Searching from left to right, the first three columns of $\mathcal{C}$ are selected since they are linearly independent. Then

$$\bar{\mathcal{C}} = [b_1, Ab_1, b_2] = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}$$

and the controllability indices are $\mu_1 = 2$ and $\mu_2 = 1$. Also, $\sigma_1 = \mu_1 = 2$ and $\sigma_2 = \mu_1 + \mu_2 = 3 = n$, and

$$\bar{\mathcal{C}}^{-1} = \begin{bmatrix} -1 & 1 & 1/2 \\ 0 & 0 & 1/2 \\ 1 & 0 & -1/2 \end{bmatrix}.$$

Notice that $q_1 = [0, 0, 1/2]$ and $q_2 = [1, 0, -1/2]$, the second and third rows of $\bar{\mathcal{C}}^{-1}$, respectively. In view of (6.53), $P = \begin{bmatrix} q_1 \\ q_1 A \\ \hline q_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1/2 \\ 0 & 1 & -1/2 \\ 1 & 0 & -1/2 \end{bmatrix}$, $P^{-1} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 2 & 0 & 0 \end{bmatrix}$, and $A_c = PAP^{-1} = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 2 & -1 & 0 \\ \hline 1 & 0 & 0 \end{array}\right]$, $B_c = PB = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \left[\begin{array}{cc} 0 & 0 \\ 1 & 1 \\ \hline 0 & 1 \end{array}\right]$.

One can also verify (6.55) quite easily. We have

$$A_c = \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 2 & -1 & 0 \\ \hline 1 & 0 & 0 \end{array}\right] = \bar{A}_c + \bar{B}_c A_m = \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \\ \hline 0 & 1 \end{array}\right] \begin{bmatrix} 2 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

and

$$B_c = \left[\begin{array}{cc} 0 & 0 \\ 1 & 1 \\ \hline 0 & 1 \end{array}\right] = \bar{B}_c B_m = \left[\begin{array}{cc} 0 & 0 \\ 1 & 0 \\ \hline 0 & 1 \end{array}\right] \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

It is interesting to note that in this example, the given pair (A, B) could have already been in controller form if B were different but A were the same. For example, consider the following three cases:

1. $A = \left[\begin{array}{c|cc} 0 & 1 & 0 \\ \hline 0 & 0 & 1 \\ 0 & 2 & -1 \end{array}\right]$, $\quad B = \left[\begin{array}{cc} 1 & \times \\ \hline 0 & 0 \\ 0 & 1 \end{array}\right]$, $\quad \mu_1 = 1, \mu_2 = 2,$
2. $A = \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 0 & 2 & -1 \end{array}\right]$, $\quad B = \left[\begin{array}{cc} 0 & 0 \\ 1 & \times \\ \hline 0 & 1 \end{array}\right]$, $\quad \mu_1 = 2, \mu_1 = 1,$

3. $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 2 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mu_1 = 3 = n.$

Note that case 3 is the single-input case (6.36).

Remarks

(i) An important result involving the controllability indices of (A, B) is the following: Given (A, B) controllable, then $(P(A + BGF)P^{-1}, PBG)$ will have the same controllability indices, within reordering, for any P, F, and G $(|P| \neq 0, |G| \neq 0)$ of appropriate dimensions. In other words, *the controllability indices are invariant under similarity and input transformations P and G, and state feedback F [or similarity transformation P and state feedback (F, G)].* (For further discussion, see [1, Subsection 3.4D].)

(ii) It is not difficult to derive explicit expressions for A_m and B_m in (6.55). Using

$$\begin{aligned} &q_i A^{k-1} b_j = 0 \quad k = 1, \ldots, \mu_j, \quad i \neq j, \\ &q_i A^{k-1} b_i = 0 \quad k = 1, \ldots, \mu_i - 1, \text{ and } q_i A^{\mu_i - 1} b_i = 1, \quad i = j, \end{aligned} \tag{6.56}$$

where $i = 1, \ldots, m$, and $j = 1, \ldots, m$, it can be shown that the m σ_1th, σ_2th, $\ldots, \sigma_m$th rows of A_c that are denoted by A_m in (6.55) are given by

$$A_m = \begin{bmatrix} q_1 A^{\mu_1} \\ \vdots \\ q_m A^{\mu_m} \end{bmatrix} P^{-1}. \tag{6.57}$$

Similarly

$$B_m = \begin{bmatrix} q_1 A^{\mu_1 - 1} \\ \vdots \\ q_m A^{\mu_m - 1} \end{bmatrix} B. \tag{6.58}$$

The matrix B_m is an upper triangular matrix with ones on the diagonal. (For details, see [1, Subsection 3.4D].)

Example 6.18. We wish to reduce $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 2 & -1 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ to controller form. Note that A and B are almost the same as in Example 6.17; however, here $\mu_1 = 1 < 2 = \mu_2$, as will be seen. We have $\mathcal{C} = [B, AB, A^2B] = [b_1, b_2, Ab_1, Ab_2, \ldots] = \begin{bmatrix} 1 & 1 & 0 & 1 & \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 2 & \end{bmatrix}$. Searching from left to right, the first three linearly independent columns are b_1, b_2, Ab_2, and $\bar{\mathcal{C}} = [b_1, b_2, Ab_2] = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$, from which we conclude that $\mu_1 = 1$, $\mu_2 = 2$, $\sigma_1 = 1$, and $\sigma_2 = 3$. We compute $\bar{\mathcal{C}}^{-1} = \begin{bmatrix} 1 & -1 & -1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{bmatrix}$. Note that $q_1 = [1, -1, -1/2]$ and $q_2 = [0, 0, 1/2]$, the first and third rows of $\bar{\mathcal{C}}^{-1}$, respectively. Then $P = \begin{bmatrix} q_1 \\ q_2 \\ q_2 A \end{bmatrix} = \begin{bmatrix} 1 & -1 & -1/2 \\ 0 & 0 & 1/2 \\ 0 & 1 & -1/2 \end{bmatrix}$, $P^{-1} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 2 & 0 \end{bmatrix}$, and

$$A_c = PAP^{-1} = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{c|cc} 0 & -1 & 0 \\ \hline 0 & 0 & 1 \\ 0 & 2 & -1 \end{array}\right],$$

$$B_c = PB = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \left[\begin{array}{cc} 1 & 0 \\ \hline 0 & 0 \\ 0 & 1 \end{array}\right].$$

It is easy to verify relations (6.57) and (6.58).

Structure Theorem—Controllable Version

The transfer function matrix $H(s)$ of the system $\dot{x} = Ax + Bu$, $y = Cx + Du$ is given by $H(s) = C(sI - A)^{-1}B + D$. If (A, B) is in *controller form* (6.54), then $H(s)$ can alternatively be characterized by the Structure Theorem stated in Theorem 6.19 below. This result is very useful in the realization of systems, which is addressed in Chapter 8 and in the study of state feedback in Chapter 9.

Let $A = A_c = \bar{A}_c + \bar{B}_c A_m$ and $B = B_c = \bar{B}_c B_m$, as in (6.55), with $|B_m| \neq 0$, and let $C = C_c$ and $D = D_c$. Define

$$\Lambda(s) \triangleq \text{diag}[s^{\mu_1}, s^{\mu_2}, \ldots, s^{\mu_m}], \tag{6.59}$$

$$S(s) \triangleq \text{block diag}([1, s, \ldots, s^{\mu_i - 1}]^T, \quad i = 1, \ldots, m). \tag{6.60}$$

Note that $S(s)$ is an $n \times m$ polynomial matrix ($n = \sum_{i=1}^m \mu_i$), i.e., a matrix with polynomials as entries. Now define the $m \times m$ polynomial matrix $D(s)$ and the $p \times m$ polynomial matrix $N(s)$ by

$$D(s) \triangleq B_m^{-1}[\Lambda(s) - A_m S(s)], N(s) \triangleq C_c S(s) + D_c D(s). \tag{6.61}$$

The following is the controllable version of the *Structure Theorem.*

Theorem 6.19. *$H(s) = N(s)D^{-1}(s)$, where $N(s)$ and $D(s)$ are defined in (6.61).*

Proof. First, note that

$$(sI - A_c)S(s) = B_c D(s). \tag{6.62}$$

To see this, we write $B_c D(s) = \bar{B}_c B_m B_m^{-1}[\Lambda(s) - A_m S(s)] = \bar{B}_c \Lambda(s) - \bar{B}_c A_m S(s)$ and $(sI - A_c)S(s) = sS(s) - (\bar{A}_c + \bar{B}_c A_m)S(s) = (sI - \bar{A}_c)S(s) - \bar{B}_c A_m S(s) = \bar{B}_c \Lambda(s) - \bar{B}_c A_m S(s)$, which proves (6.62). Now $H(s) = C_c(sI - A_c)^{-1}B_c + D_c = C_c S(s) D^{-1}(s) + D_c = [C_c S(s) + D_c D(s)]D^{-1}(s) = ND^{-1}$. ■

Example 6.20. Let $A_c = \begin{bmatrix} 0 & 1 & 0 \\ 2 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $B_c = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$, as in Example 6.17. Here $\mu_1 = 2, \mu_2 = 1$ and $A_m = \begin{bmatrix} 2 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$, $B_m = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Then $\Lambda(s) = \begin{bmatrix} s^2 & 0 \\ 0 & s \end{bmatrix}$, $S(s) = \begin{bmatrix} 1 & 0 \\ s & 0 \\ 0 & 1 \end{bmatrix}$ and

$$\begin{aligned} D(s) = B_m^{-1}[\Lambda(s) - A_m S(s)] &= \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \left[\begin{bmatrix} s^2 & 0 \\ 0 & s \end{bmatrix} - \begin{bmatrix} -s+2 & 0 \\ 1 & 0 \end{bmatrix} \right] \\ &= \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} s^2+s-2 & 0 \\ -1 & s \end{bmatrix} = \begin{bmatrix} s^2+s-1 & -s \\ -1 & s \end{bmatrix}. \end{aligned}$$

Now $C_c = [0,\ 1,\ 1]$, and $D_c = [0,\ 0]$,

$$N(s) = C_c S(s) + D_c D(s) = [s,\ 1],$$

and

$$\begin{aligned} H(s) = [s, 1] \begin{bmatrix} s^2+s-1 & -s \\ -1 & s \end{bmatrix}^{-1} &= [s, 1] \begin{bmatrix} s & s \\ 1 & s^2+s-1 \end{bmatrix} \frac{1}{s(s^2+s-2)} \\ &= \frac{1}{s(s^2+s-2)} [s^2+1, 2s^2+s-1] \\ &= C_c(sI - A_c)^{-1}B_c + D_c. \end{aligned}$$

Example 6.21. Let $A_c = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 1 & -2 \end{bmatrix}$, $B_c = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, $C_c = [0,\ 1,\ 0]$, and $D_c = 0$ (see Example 6.13). In the present case, we have $A_m = [2, 1, -2]$, $B_m = 1$, $\Lambda(s) = s^3$, $S(s) = [1, s, s^2]^T$, and

$$D(s) = 1 \cdot [s^3 - [2,\ 1,\ -2][1,\ s,\ s^2]^T] = s^3 + 2s^2 - s - 2, \quad N(s) = s.$$

Then

$$H(s) = N(s)D^{-1}(s) = s/(s^3 + 2s^2 - s - 2) = C_c(sI - A_c)^{-1}B_c + D_c.$$

6.4.2 Observer Forms

Consider the system $\dot{x} = Ax + Bu$, $y = Cx + Du$ given in (6.1) and assume that (A, C) is observable; i.e., $\operatorname{rank} \mathcal{O} = n$, where

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}. \tag{6.63}$$

Also, assume that the $p \times n$ matrix C has a full row rank p; i.e.,

$$\operatorname{rank} C = p \leq n. \tag{6.64}$$

It is of interest to determine a transformation matrix P so that the equivalent system representation $\{A_o, B_o, C_o, D_o\}$ with

$$A_o = PAP^{-1}, \quad B_o = PB, \quad C_o = CP^{-1}, \quad D_o = D \tag{6.65}$$

will have (A_o, C_o) in an observer form (defined below). As will become clear in the following discussion, these forms are dual to the controller forms previously discussed and can be derived by taking advantage of this fact. In particular, let $\widetilde{A} \triangleq A^T$, $\widetilde{B} \triangleq C^T$ [$(\widetilde{A}, \widetilde{B})$ is controllable], and determine a nonsingular transformation $\widetilde{P}$ so that $\widetilde{A}_c = \widetilde{P}\widetilde{A}\widetilde{P}^{-1}$, $\widetilde{B}_c = \widetilde{P}\widetilde{B}$ are in controller form given in (6.54). Then $A_o = \widetilde{A}_c^T$ and $C_o = \widetilde{B}_c^T$ is in observer form.

It will be demonstrated in the following discussion how to obtain observer forms directly, in a way that parallels the approach described for controller forms. This is done for the sake of completeness and to define the observability indices. The approach of using duality just given can be used in each case to verify the results.

We first note that if $\operatorname{rank} C = r < p$, an approach analogous to the case when $\operatorname{rank} B < m$ can be followed, as in Subsection 6.4.1. The fact that the rows of C are not linearly independent means that the same information can be extracted from only r outputs, and therefore, the choice for the outputs should perhaps be reconsidered. Now if (A, C) is unobservable, one may use two steps to first isolate the observable part and then reduce it to the observer form, in an analogous way to the uncontrollable case previously given.

Single-Output Case ($p = 1$)

Let

$$P^{-1} = Q \triangleq [\tilde{q}, A\tilde{q}, \ldots, A^{n-1}\tilde{q}], \tag{6.66}$$

where $\tilde{q}$ is the nth column in $\mathcal{O}^{-1}$. Then

$$A_0 = \begin{bmatrix} 0 \cdots 0 & -\alpha_0 \\ 1 \cdots 0 & -\alpha_1 \\ \vdots \ddots \vdots & \vdots \\ 0 \cdots 1 & -\alpha_{n-1} \end{bmatrix}, \quad C_o = [0, \ldots, 0, 1], \tag{6.67}$$

where the α_i denote the coefficients of the characteristic polynomial $\alpha(s) \triangleq \det(sI - A) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s + \alpha_0$. Here $A_o = PAP^{-1} = Q^{-1}AQ, C_o = CP^{-1} = CQ$, and the desired result can be established by using a proof that is completely analogous to the proof in determining the (dual) controller form presented in Subsection 6.4.1. Note that $B_o = PB$ does not have any particular structure. The representation $\{A_o, B_o, C_o, D_o\}$ will be referred to as the *observer form* of the system.

Reversing the order of columns in P^{-1} given in (6.66) or selecting P to be exactly $\mathcal{O}$, or to be equal to the matrix obtained after the order of the columns in $\mathcal{O}$ has been reversed, leads to alternative observer forms in a manner analogous to the controller form case.

Example 6.22. Let $A = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$ and $C = [1, -1, 1]$. To derive the observer form (6.67), we could use duality, by defining $\widetilde{A} = A^T, \widetilde{B} = C^T$, and deriving the controller form of $\widetilde{A}, \widetilde{B}$, i.e., by following the procedure outlined above. We note that the $\widetilde{A}, \widetilde{B}$ are exactly the matrices given in Examples 6.13 and 6.14. As an alternative approach, the observer form is now derived directly. In particular, we have

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ -1 & -1 & -2 \\ 1 & -1 & 4 \end{bmatrix}, \mathcal{O}^{-1} = \begin{bmatrix} 1 & -1/2 & -1/2 \\ -1/3 & -1/2 & -1/6 \\ -1/3 & 0 & 1/3 \end{bmatrix},$$

and in view of (6.66),

$$Q = P^{-1} = [\tilde{q}, A\tilde{q}, A^2\tilde{q}] = \begin{bmatrix} -1/2 & 1/2 & -1/2 \\ -1/6 & -1/6 & -1/6 \\ 1/3 & -2/3 & 4/3 \end{bmatrix}.$$

Note that $\tilde{q} = [-1/2, -1/6, 1/3]^T$, the last column of $\mathcal{O}^{-1}$. Then

$$A_o = Q^{-1}AQ = \begin{bmatrix} 0 & 0 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & -2 \end{bmatrix}, \text{ and } C_o = CQ = [0, 0, 1],$$

where $|sI - A| = s^3 + 2s - s - 2 = s^3 + \alpha_2 s^2 + \alpha_1 s + \alpha_0$. Hence, $QA_o = \begin{bmatrix} 1/2 & -1/2 & 1/2 \\ -1/6 & -1/6 & -1/6 \\ -2/3 & 4/3 & -8/3 \end{bmatrix} = AQ$.

Multi-Output Case ($p > 1$)

Consider

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_p \\ c_1 A \\ \vdots \\ c_p A \\ \vdots \\ c_1 A^{n-1} \\ \vdots \\ c_p A^{n-1} \end{bmatrix}, \tag{6.68}$$

where $c_1, \dots, c_p$ denote the p rows of C, and select the first n linearly independent rows in $\mathcal{O}$, moving from the top to bottom (rank $\mathcal{O} = n$). Next, reorder the selected rows by first taking all rows involving c_1, then c_2, etc., to obtain

$$\bar{\mathcal{O}} \triangleq \begin{bmatrix} c_1 \\ c_1 A \\ \vdots \\ c_1 A^{\nu_1 - 1} \\ \vdots \\ c_p \\ \vdots \\ c_p A^{\nu_p - 1} \end{bmatrix}, \tag{6.69}$$

an $n \times n$ matrix. The integer ν_i denotes the number of rows involving c_i in the set of the first n linearly independent rows found in $\mathcal{O}$ when moving from top to bottom.

Definition 6.23. *The p integers ν_i, $i = 1, \dots, p$, are the* observability indices of the system, *and $\nu \triangleq \max \nu_i$ is called* the observability index of the system. *Note that*

$$\sum_{i=1}^{p} \nu_i = n \quad \textit{and} \quad p\nu \geq n. \tag{6.70}$$

■

When $\operatorname{rank} C = p$, then $\nu_i \geq 1$. Now define

$$\tilde{\sigma}_k \triangleq \sum_{i=1}^{k} \nu_i \quad k = 1, \dots, p; \tag{6.71}$$

i.e., $\tilde{\sigma}_1 = \nu_1, \tilde{\sigma}_2 = \nu_1 + \nu_2, \ldots, \tilde{\sigma}_p = \nu_1 + \cdots + \nu_p = n$. Consider $\bar{\mathcal{O}}^{-1}$ and let $\tilde{q}_k \in R^n$, $k = 1, \ldots, p$, represent its $\tilde{\sigma}_k$th column; i.e.,

$$\bar{\mathcal{O}}^{-1} = [\times \cdots \times \tilde{q}_1 | \times \cdots \times \tilde{q}_2 | \cdots | \times \cdots \times \tilde{q}_p]. \tag{6.72}$$

Define

$$P^{-1} = Q = [\tilde{q}_1, \ldots, A^{\nu_1 - 1}\tilde{q}_1, \ldots, \tilde{q}_p, \ldots, A^{\nu_p - 1}\tilde{q}_p]. \tag{6.73}$$

Then $A_o = PAP^{-1} = Q^{-1}AQ$ and $C_o = CP^{-1} = CQ$ are given by

$$A_o = [A_{ij}], \quad i, j = 1, \ldots, p,$$

$$A_{ii} = \begin{bmatrix} 0 \cdots 0 & \times \\ I_{\nu_i - 1} & \vdots \\ & \times \end{bmatrix} \in R^{\nu_i \times \nu_i}, \quad i = j, A_{ij} = \begin{bmatrix} 0 \cdots & 0 & \times \\ \vdots & \vdots & \vdots \\ 0 \cdots & 0 & \times \end{bmatrix} \in R^{\nu_i \times \nu_j}, \quad i \neq j,$$

and

$$C_o = [C_1, C_2, \ldots, C_p], C_i = \begin{bmatrix} 0 \cdots & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 \cdots & 0 & 0 \\ 0 \cdots & 0 & 1 \\ 0 \cdots & 0 & \times \\ \vdots & \vdots & \vdots \\ 0 \cdots & 0 & \times \end{bmatrix} \in R^{p \times \nu_i}, \tag{6.74}$$

where the 1 in the last column of C_i occurs at the ith row location ($i = 1, \ldots, p$) and $\times$ denotes nonfixed entries. Note that the matrix $B_o = PB = Q^{-1}B$ does not have any particular structure. Equation (6.74) is a very useful form (in the observer problem) and shall be referred to as the *observer form* of the system.

Analogous to (6.55), we express A_o and C_o as

$$A_o = \bar{A}_o + A_p\bar{C}_o, \quad C_o = C_p\bar{C}_o, \tag{6.75}$$

where $\bar{A}_o =$ block diag$[A_1, A_2, \ldots, A_p]$ with $A_i = \begin{bmatrix} 0 \cdots & 0 \\ I_{\nu_i - 1} & \vdots \\ & 0 \end{bmatrix} \in R^{\nu_i \times \nu_i}, \bar{C}_o =$ block diag$([0, \ldots, 0, 1]^T \in R^{\nu_i}$, $i = 1, \ldots, p)$, and $A_p \in R^{n \times p}$, and $C_p \in R^{p \times p}$ are appropriate matrices ($\sum_{i=1}^{p} \nu_i = n$). Note that $\bar{A}_o, \bar{C}_o$ are completely determined by the p observability indices ν_i, $i = 1, \ldots, p$, and A_p and C_p contain this information in the $\tilde{\sigma}_1$th$, \ldots, \tilde{\sigma}_p$th columns of A_o and in the same columns of C_o, respectively.

Example 6.24. Given $A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 2 \\ 0 & 1 & -1 \end{bmatrix}$ and $C = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$, we wish to reduce these to observer form. This can be accomplished using duality, i.e., by first reducing $\widetilde{A} \triangleq A^T, \widetilde{B} \triangleq C^T$ to controller form. Note that $\widetilde{A}, \widetilde{B}$ are the matrices used in Example 6.17, and therefore, the desired answer is easily obtained. Presently, we shall follow the direct algorithm described above. We have

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 2 \\ 1 & 0 & 2 \\ 0 & 2 & -2 \\ 0 & 2 & -2 \end{bmatrix}.$$

Searching from top to bottom, the first three linearly independent rows are c_1, c_2, c_1A, and

$$\bar{\mathcal{O}} = \begin{bmatrix} c_1 \\ c_1A \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 2 \\ 1 & 1 & 0 \end{bmatrix}.$$

Note that the observability indices are $\nu_1 = 2, \nu_2 = 1$ and $\tilde{\sigma}_1 = 2, \tilde{\sigma}_2 = 3$. We compute

$$\bar{\mathcal{O}}^{-1} = \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & 0 \\ 1/2 & 1/2 & -1/2 \end{bmatrix} = \begin{bmatrix} \times & 0 & 1 \\ \times & 0 & 0 \\ \times & 1/2 & -1/2 \end{bmatrix}.$$

Then, $Q = [\tilde{q}_1, A\tilde{q}_1, \tilde{q}_2] = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1/2 & -1/2 & -1/2 \end{bmatrix}$ and $Q^{-1} = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$. Therefore,

$$A_o = Q^{-1}AQ = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] = \left[\begin{array}{cc|c} 0 & 2 & 1 \\ 1 & -1 & 0 \\ \hline 0 & 0 & 0 \end{array}\right], C_o = CQ = [C_1 \vdots C_2] = \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 1 & 1 \end{array}\right].$$

We can also verify (6.47), namely

$$A_o = \left[\begin{array}{cc|c} 0 & 2 & 1 \\ 1 & -1 & 0 \\ \hline 0 & 0 & 0 \end{array}\right] = \bar{A}_o + A_p\bar{C}_o = \left[\begin{array}{cc|c} 0 & 0 & 0 \\ 1 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right] + \begin{bmatrix} 2 & 1 \\ -1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and

$$C_o = \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 1 & 1 \end{array}\right] = C_p\bar{C}_o = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Structure Theorem—Observable Version

The transfer function matrix $H(s)$ of system $\dot{x} = Ax + Bu$, $y = Cx + Du$ is given by $H(s) = C(sI - A)^{-1}B + D$. If (A, C) is in the *observer form*, given in (6.74), then $H(s)$ can alternatively be characterized by the Structure

Theorem stated in Theorem 6.25 below. This result will be very useful in the realization of systems, addressed in Chapter 8 and also in the study of observers in Chapter 9.

Let $A = A_o = \bar{A}_o + A_p\bar{C}_o$ and $C = C_o = C_p\bar{C}_o$ as in (6.75) with $|C_p| \neq 0$; let $B = B_o$ and $D = D_o$, and define

$$\widetilde{\Lambda}(s) \triangleq \text{diag}[s^{\nu_1}, s^{\nu_2}, \dots, s^{\nu_p}], \widetilde{S}(s) \triangleq \text{block diag}([1, s, \dots, s^{\nu_i - 1}], i = 1, \dots, p). \tag{6.76}$$

Note that $\widetilde{S}(s)$ is a $p \times n$ polynomial matrix, where $n = \sum_{i=1}^{p} \nu_i$. Now define the $p \times p$ polynomial matrix $\widetilde{D}(s)$ and the $p \times m$ polynomial matrix $\widetilde{N}(s)$ as

$$\widetilde{D}(s) \triangleq [\widetilde{\Lambda}(s) - \widetilde{S}(s)A_p]C_p^{-1}, \quad \widetilde{N}(s) \triangleq \widetilde{S}(s)B_o + \widetilde{D}(s)D_o. \tag{6.77}$$

The following result is the observable version of the *Structure Theorem*. It is the dual of Theorem 6.19 and can therefore be proved using duality arguments. The proof given is direct.

Theorem 6.25. $H(s) = \widetilde{D}^{-1}(s)\widetilde{N}(s)$, *where* $\widetilde{N}(s), \widetilde{D}(s)$ *are defined in (6.77).*

Proof. First we note that

$$\widetilde{D}(s)C_o = \widetilde{S}(s)(sI - A_o). \tag{6.78}$$

To see this, write $\widetilde{D}(s)C_o = [\widetilde{\Lambda}(s) - \widetilde{S}(s)A_p]C_p^{-1}C_p\bar{C}_o = \widetilde{\Lambda}(s)\bar{C}_o - \widetilde{S}(s)A_p\bar{C}_o$, and also, $\widetilde{S}(s)(sI - A_o) = \widetilde{S}(s)s - \widetilde{S}(s)(\bar{A}_o + A_p\bar{C}_o) = \widetilde{S}(s)(sI - \bar{A}_o) - \widetilde{S}(s)A_p\bar{C}_o = \widetilde{\Lambda}(s)\bar{C}_o - \widetilde{S}(s)A_p\bar{C}_o$, which proves (6.78). We now obtain $H(s) = C_o(sI - A_o)^{-1}B_o + D_o = \widetilde{D}^{-1}(s)\widetilde{S}(s)B_o + D_o = \widetilde{D}^{-1}(s)[\widetilde{S}(s)B_o + \widetilde{D}(s)D_o] = \widetilde{D}^{-1}(s)\widetilde{N}(s)$. ■

Example 6.26. Consider $A_o = \begin{bmatrix} 0 & 2 & 1 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $C_o = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ of Example 6.24. Here $\nu_1 = 2, \nu_2 = 1$, $\widetilde{\Lambda}(s) = \begin{bmatrix} s^2 & 0 \\ 0 & s \end{bmatrix}$, and $\widetilde{S}(s) = \begin{bmatrix} 1 & s & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Then $\widetilde{D}(s) = [\widetilde{\Lambda}(s) - \widetilde{S}(s)A_p]C_p^{-1} = \left[\begin{bmatrix} s^2 & 0 \\ 0 & s \end{bmatrix} - \begin{bmatrix} 1 & s & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & 0 \\ 0 & 0 \end{bmatrix}\right] \cdot \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1} =$ $\left[\begin{bmatrix} s^2 & 0 \\ 0 & s \end{bmatrix} - \begin{bmatrix} -s+2 & 1 \\ 0 & 0 \end{bmatrix}\right] \cdot \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} s^2 + s - 2, & -1 \\ 0 & s \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} s^2 + s - 1 & -1 \\ -s & s \end{bmatrix}$.

Now if $B_o = [0,\ 1,\ 1]^T, D_o = 0$, and $\widetilde{N}(s) = \widetilde{S}(s)B_o + \widetilde{D}(s)D_o = [s, 1]^T$, then $H(s) = \widetilde{D}^{-1}(s)\widetilde{N}(s) = \frac{1}{s(s^2+s-2)}[s^2+1, 2s^2+s-1]^T = C_o(sI - A_o)^{-1}B_o + D_o$.

6.5 Summary and Highlights

- The standard form for uncontrollable systems is

$$\widehat{A} = Q^{-1}AQ = \begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix}, \quad \widehat{B} = Q^{-1}B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \tag{6.6}$$

where $A_1 \in R^{n_r \times n_r}$, $B_1 \in R^{n_r \times m}$, and (A_1, B_1) is controllable. $n_r < n$ is the rank of the controllability matrix $\mathcal{C} = [B, AB, \ldots, A^{n-1}B]$; i.e.,

$$\operatorname{rank} \mathcal{C} = n_r.$$

- The standard form for unobservable systems is

$$\widehat{A} = Q^{-1}AQ = \begin{bmatrix} A_1 & 0 \\ A_{21} & A_2 \end{bmatrix}, \quad \widehat{C} = CQ = \begin{bmatrix} C_1 \\ 0 \end{bmatrix}, \tag{6.15}$$

where $A_1 \in R^{n_o \times n_o}$, $C_1 \in R^{p \times n_o}$, and (A_1, C_1) is observable. $n_o < n$ is the rank of the observability matrix

$$\mathcal{O} = \begin{bmatrix} \widehat{C} \\ \widehat{C}\widehat{A} \\ \vdots \\ \widehat{C}\widehat{A}^{n-1} \end{bmatrix};$$

i.e.,

$$\operatorname{rank} \mathcal{O} = n_o.$$

- Kalman's Decomposition Theorem.

$$\widehat{A} = Q^{-1}AQ = \begin{bmatrix} A_{11} & 0 & A_{13} & 0 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ 0 & 0 & A_{33} & 0 \\ 0 & 0 & A_{43} & A_{44} \end{bmatrix}, \quad \widehat{B} = Q^{-1}B = \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{bmatrix}, \quad \widehat{C} = CQ = [C_1,\ 0,\ C_3,\ 0], \tag{6.20}$$

where (A_{11}, B_1, C_1) is controllable and observable.
- λ_i is an uncontrollable eigenvalue if and only if

$$\hat{v}_i[\lambda_i I - A, B] = 0, \tag{6.23}$$

where $\hat{v}_i$ is the corresponding (left) eigenvector.
- λ_i is an unobservable eigenvalue if and only if

$$\begin{bmatrix} \lambda_i I - A \\ C \end{bmatrix} v_i = 0, \tag{6.24}$$

where v_i is the corresponding (right) eigenvector.

Controller Forms (for Controllable Systems)

- $m = 1$ case.

$$A_c = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ -\alpha_0 & -\alpha_1 & \cdots & -\alpha_{n-1} \end{bmatrix}, \quad B_c = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \tag{6.36}$$

where

$$\alpha(s) \triangleq \det(sI - A) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_1 s + \alpha_0. \tag{6.37}$$

- $m > 1$ case.

$$A_c = [A_{ij}], \qquad i, j = 1, \ldots, m,$$

$$A_{ii} = \begin{bmatrix} 0 & \\ \vdots & I_{\mu_i - 1} \\ 0 & \\ \times & \times \cdots \times \end{bmatrix} \in R^{\mu_i \times \mu_i},\ i = j,\ A_{ij} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \cdots & 0 \\ \times & \times \cdots & \times \end{bmatrix} \in R^{\mu_i \times \mu_j},\ i \neq j,$$

and

$$B_c = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{bmatrix}, \qquad B_i = \begin{bmatrix} 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & 1 & \times \cdots & \times \end{bmatrix} \in R^{\mu_i \times m}. \tag{6.54}$$

An example for $n = 4$, $m = 2$ and $\mu_1 = 2, \mu_2 = 2$ is

$$A_c = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ \times & \times & \times & \times \\ \hline 0 & 0 & 0 & 1 \\ \times & \times & \times & \times \end{array}\right], \quad B_c = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \left[\begin{array}{cc} 0 & 0 \\ 1 & \times \\ \hline 0 & 0 \\ 0 & 1 \end{array}\right].$$

- $$A_c = \bar{A}_c + \bar{B}_c A_m, \quad B_c = \bar{B}_c B_m. \tag{6.55}$$

- *Structure theorem—controllable version*
$H(s) = N(s)D^{-1}(s)$, where

$$D(s) = B_m^{-1}[\Lambda(s) - A_m S(s)], N(s) = C_c S(s) + D_c D(s). \tag{6.61}$$

Note that

$$(sI - A_c)S(s) = B_c D(s). \tag{6.62}$$

Observer Forms (for Observable Systems)

- $p = 1$ case.

$$A_0 = \begin{bmatrix} 0 \cdots 0 & -\alpha_0 \\ 1 \cdots 0 & -\alpha_1 \\ \vdots \ddots \vdots & \vdots \\ 0 \cdots 1 & -\alpha_{n-1} \end{bmatrix}, \quad C_o = [0, \ldots, 0, 1]. \tag{6.67}$$

- $p > 1$.

$$A_o = [A_{ij}], \quad i, j = 1, \ldots, p,$$

$$A_{ii} = \begin{bmatrix} 0 \cdots 0 & \times \\ I_{\nu_i - 1} & \vdots \\ & \times \end{bmatrix} \in R^{\nu_i \times \nu_i},\ i = j,\ A_{ij} = \begin{bmatrix} 0 \cdots 0 & \times \\ \vdots \quad \vdots & \vdots \\ 0 \cdots 0 & \times \end{bmatrix} \in R^{\nu_i \times \nu_j},\ i \neq j,$$

and

$$C_o = [C_1, C_2, \ldots, C_p], C_i = \begin{bmatrix} 0 \cdots 0 & 0 \\ \vdots \quad \vdots & \vdots \\ 0 \cdots 0 & 0 \\ 0 \cdots 0 & 1 \\ 0 \cdots 0 & \times \\ \vdots \quad \vdots & \vdots \\ 0 \cdots 0 & \times \end{bmatrix} \in R^{p \times \nu_i}, \tag{6.74}$$

If (A_c, B_c) is in controller form, $(A_o = A_c^T, C_o = B_c^T)$ will be in observer form.

- $$A_o = \bar{A}_o + A_p \bar{C}_o, \quad C_o = C_p \bar{C}_o. \tag{6.75}$$

- *Structure theorem—observable version*
  $H(s) = \widetilde{D}^{-1}(s)\widetilde{N}(s)$, where

$$\widetilde{D}(s) = [\widetilde{A}(s) - \widetilde{S}(s)A_p]C_p^{-1}, \quad \widetilde{N}(s) = \widetilde{S}(s)B_o + \widetilde{D}(s)D_o. \tag{6.77}$$

Note that

$$\widetilde{D}(s)C_o = \widetilde{S}(s)(sI - A_o). \tag{6.78}$$

6.6 Notes

Special state-space forms for controllable and observable systems obtained by similarity transformations are discussed at length in Kailath [5]. Wolovich [13] discusses the algorithms for controller and observer forms and introduces the Structure Theorems. The controller form is based on results by Luenberger [9] (see also Popov [10]). A detailed derivation of the controller form can also be found in Rugh [12].

Original sources for the Canonical Structure Theorem include Kalman [6] and Gilbert [3].

The eigenvector and rank tests for controllability and observability are called PBH tests in Kailath [5]. Original sources for these include Popov [10], Belevich [2], and Hautus [4]. Consult also Rosenbrock [11], and for the case when A can be diagonalized via a similarity transformation, see Gilbert [3]. Note that in the eigenvalue/eigenvector tests presented herein the uncontrollable (unobservable) eigenvalues are also explicitly identified, which represents a modification of the above original results.

The fact that the controllability indices appear in the work of Kronecker was recognized by Rosenbrock [11] and Kalman [8].

For an extensive introductory discussion and a formal definition of canonical forms, see Kailath [5].

References

1. P.J. Antsaklis and A.N. Michel, *Linear Systems*, Birkhäuser, Boston, MA, 2006.
2. V. Belevich, *Classical Network Theory*, Holden-Day, San Francisco, CA, 1968.
3. E. Gilbert, "Controllability and observability in multivariable control systems," *SIAM J. Control*, Vol. 1, pp. 128–151, 1963.
4. M.L.J. Hautus, "Controllability and observability conditions of linear autonomous systems," *Proc. Koninklijke Akademie van Wetenschappen, Serie A*, Vol. 72, pp. 443–448, 1969.
5. T. Kailath, *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.
6. R.E. Kalman, "Mathematical descriptions of linear systems," *SIAM J. Control*, Vol. 1, pp. 152–192, 1963.
7. R.E. Kalman, "On the computation of the reachable/observable canonical form," *SIAM J. Control Optimization*, Vol. 20, no. 2, pp. 258–260, 1982.
8. R.E. Kalman, "Kronecker invariants and feedback," in *Ordinary Differential Equations*, L. Weiss, ed., pp. 459–471, Academic Press, New York, NY, 1972.
9. D.G. Luenberger, "Canonical forms for linear multivariable systems," *IEEE Trans. Auto. Control*, Vol. 12, pp. 290–293, 1967.
10. V.M. Popov, "Invariant description of linear, time-invariant controllable systems," *SIAM J. Control Optimization*, Vol. 10, No. 2, pp. 252–264, 1972.
11. H.H. Rosenbrock, *State-Space and Multivariable Theory*, Wiley, New York, NY, 1970.
12. W.J. Rugh, *Linear System Theory*, Second Ed., Prentice-Hall, Englewood Cliffs, NJ, 1996.
13. W.A. Wolovich, *Linear Multivariable Systems*, Springer-Verlag, New York, NY, 1974.

Exercises

6.1. Write software programs to implement the algorithms of Section 6.2. In particular:

(a) Given the pair (A, B), where $A \in R^{n\times n}, B \in R^{n\times m}$ with

$$\text{rank}[B, AB, \ldots, A^{n-1}B] = n_r < n,$$

reduce this pair to the standard uncontrollable form

$$\widehat{A} = PAP^{-1} = \begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix}, \widehat{B} = PB = \begin{bmatrix} B_1 \\ 0 \end{bmatrix},$$

where (A_1, B_1) is controllable and $A_1 \in R^{n_r\times n_r}, B_1 \in R^{n_r\times m}$.

(b) Given the controllable pair (A, B), where $A \in R^{n\times n}, B \in R^{n\times m}$ with $\text{rank}\, B = m$, reduce this pair to the controller form $A_c = PAP^{-1}, B_c = PB$.

6.2. Determine the uncontrollable modes of each pair (A, B) given below by

(a) Reducing (A, B), using a similarity transformation.
(b) Using eigenvalue/eigenvector criteria:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

6.3. Reduce the pair

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

into controller form $A_c = PAP^{-1}, B_c = PB$. What is the similarity transformation matrix in this case? What are the controllability indices?

6.4. Consider

$$A_c = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ -\alpha_0 & -\alpha_1 & \cdots & -\alpha_{n-1} \end{bmatrix}, \quad B_c = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

Show that

$$\mathcal{C} = [B_c, A_cB_c, \ldots, A_c^{n-1}B_c] = \begin{bmatrix} 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & c_1 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 1 & \cdots & c_{n-3} \\ 0 & 1 & c_1 & \cdots & c_{n-2} \\ 1 & c_1 & c_2 & \cdots & c_{n-1} \end{bmatrix},$$

where $c_k = -\sum_{i=0}^{k-1} \alpha_{n-i-1} c_{k-i-1}$, $k = 1, \ldots, n-1$, with $c_0 = 1$. Also, show that

$$\mathcal{C}^{-1} = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{n-1} & 1 \\ \alpha_2 & \alpha_3 & \cdots & 1 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ \alpha_{n-1} & 1 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

6.5. Show that the matrices $A_c = PAP^{-1}, B_c = PB$ are as follows:

(a) Given by (6.41) if P is given by (6.42).
(b) Given by (6.44) if $Q(= P^{-1})$ is given by (6.43).
(c) Given by (6.46) if $Q(= P^{-1})$ is given by (6.45).

6.6. Consider the pair (A, b), where $A \in R^{n\times n}, b \in R^n$. Show that if more than one linearly independent eigenvector can be associated with a single eigenvalue, then (A, b) is uncontrollable. *Hint:* Use the eigenvector test. Let $\hat{v}_1, \hat{v}_2$ be linearly independent left eigenvectors associated with eigenvalue $\lambda_1 = \lambda_2 = \lambda$. Notice that if $\hat{v}_1 b = \alpha_1$ and $\hat{v}_2 b = \alpha_2$, then $(\alpha_1 \hat{v}_1 - \alpha_1 \hat{v}_2)b = 0$.

6.7. Show that if (A, B) is controllable, where $A \in R^{n\times n}$, and $B \in R^{n\times m}$, and $\operatorname{rank} B = m$, then $\operatorname{rank} A \geq n - m$.

6.8. Given $A \in R^{n\times n}$, and $B \in R^{n\times m}$, let $\operatorname{rank} \mathcal{C} = n$, where $\mathcal{C} = [B, AB, \ldots, A^{n-1}B]$. Consider $\widehat{A} \in R^{n\times n}, \widehat{B} \in R^{n\times m}$ with $\operatorname{rank} \widehat{\mathcal{C}} = n$, where $\widehat{\mathcal{C}} = [\widehat{B}, \widehat{A}\widehat{B}, \ldots, \widehat{A}^{n-1}\widehat{B}]$, and assume that $P \in R^{n\times n}$ with $\det P \neq 0$ exists such that

$$P[\mathcal{C}, A^n B] = [\widehat{\mathcal{C}}, \widehat{A}^n \widehat{B}].$$

Show that $\widehat{B} = PB$ and $\widehat{A} = PAP^{-1}$. *Hint:* Show that $(PA - \widehat{A}P)\mathcal{C} = 0$.

6.9. Let $A = \bar{A}_c + \bar{B}_c A_m$ and $B = \bar{B}_c B_m$, where the $\bar{A}_c, \bar{B}_c$ are as in (6.55) with $A_m \in R^{m\times n}, B_m \in R^{m\times m}$, and $|B_m| \neq 0$. Show that (A, B) is controllable with controllability indices μ_i. *Hint:* Use the eigenvalue test to show that (A, B) is controllable. Use state feedback to simplify (A, B) (see Exercise 6.11), and show that the μ_i are the controllability indices.

6.10. Show that the controllability indices of the state equation $\dot{x} = Ax + BGv$, where $|G| \neq 0$ and (A, B) is controllable, with $A \in R^{n\times n}$, $B \in R^{n\times m}$, are the same as the controllability indices of $\dot{x} = Ax + Bu$, within reordering. *Hint:* Write $\bar{\mathcal{C}}_k = [BG, ABG, \ldots, A^{k-1}BG] = [B, AB, \ldots, A^{k-1}B] \cdot [\text{block diag}\, G] = \mathcal{C}_k \cdot [\text{block diag}\, G]$ and show that the number of linearly dependent columns in $A^k BG$ that occur while searching from left to right in $\bar{\mathcal{C}}_n$ is the same as the corresponding number in $\mathcal{C}_n$.

6.11. Consider the state equation $\dot{x} = Ax + Bu$, where $A \in R^{n\times n}, B \in R^{n\times m}$ with (A, B) controllable. Let the linear state-feedback control law be $u = Fx + Gv, F \in R^{m\times n}, G \in R^{m\times m}$ with $|G| \neq 0$. Show that

(a) $(A + BF, BG)$ is controllable.
(b) The controllability indices of $(A+BF, B)$ are identical to those of (A, B).
(c) The controllability indices of $(A+BF, BG)$ are equal to the controllability indices of (A, B) within reordering. *Hint:* Use the eigenvalue test to show (a). To show (b), use the controller forms in Section 6.4.

6.12. For the system $\dot{x} = Ax + Bu, y = Cx$, consider the corresponding sampled-data system $\bar{x}(k+1) = \bar{A}\bar{x}(k) + \bar{B}\bar{u}(k), \bar{y}(k) = \bar{C}\bar{x}(k)$, where

$$\bar{A} = e^{AT}, \bar{B} = [\int_0^T e^{A\tau} d\tau]B, \quad \text{and} \quad \bar{C} = C.$$

(a) Let the continuous-time system $\{A, B, C\}$ be controllable (observable), and assume it is a SISO system. Show that $\{\bar{A}, \bar{B}, \bar{C}\}$ is controllable (observable) if and only if the sampling period T is such that

$$Im\ (\lambda_i - \lambda_j) \neq \frac{2\pi k}{T}, \text{ where } k = \pm 1, \pm 2, \ldots \text{ whenever } Re\ (\lambda_i - \lambda_j) = 0,$$

where $\{\lambda_i\}$ are the eigenvalues of A. *Hint:* Use the PBH test.
(b) Apply the results of (a) to the double integrator (Example 3.33 in Chapter 3), where $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and $C = [1, 0]$, and also to $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $C = [1, 0]$. Determine the values of T that preserve controllability (observability).

6.13. (**Spring mass system**) Consider the spring mass given in Exercise 3.37.

(a) Is the system controllable from $[f_1, f_2]^T$? If yes, reduce (A, B) to controller form.
(b) Is the system controllable from input f_1 only? Is it controllable from f_2 only? Discuss your answers.
(c) Let $y = Cx$ with $C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$. Is the system observable from y? If yes, reduce (A, C) to observer form.