CHAPTER 6

CONTROLLABILITY, OBSERVABILITY, AND DISTANCE TO UNCONTROLLABILITY

Topics covered

- Controllability and Observability Criteria
- Controller- and Observer-Companion Forms
- Kalman Decomposition
- Controller and Observer Hessenberg Forms
- Distance to Uncontrollability

6.1 INTRODUCTION

This chapter deals with discussions on the two most fundamental notions, **controllability** and **observability**, and related concepts. The well-known criteria of controllability and observability are stated and proved in Theorem 6.2.1.

These theoretically important criteria, unfortunately, do not yield numerically effective tests of controllability. This is demonstrated by means of some examples and discussions in **Section 6.6**. **Numerically effective tests**, based on reduction of the pairs (A, B) and (A, C), respectively, to the **controller-Hessenberg** and **observer-Hessenberg** pairs, achieved by means of orthogonal similarly, are described in **Sections 6.7 and 6.8**.

Controllability and observability are generic concepts. What is more important in practice is to know when a controllable system is close to an uncontrollable one. To this end, a measure of the distance to uncontrollability is introduced in **Section 6.9** and a characterization **(Theorem 6.9.1)** in terms of the minimum singular value of a certain matrix is stated and proved. Finally, two algorithms (**Algorithms 6.9.1 and 6.9.2**) are described to measure the distance to uncontrollability.

The chapter concludes with a brief discussion (**Section 6.10**) on the relationship between the distance to uncontrollability and the singular values of the controllability matrix. *The important message here is that the singular values of the controllability matrix as such cannot be used to make a prediction of how close the system is to an uncontrollable system.* It is the largest gap between two singular values that should be considered.

Reader's Guide for Chapter 6

The readers having knowledge of basic concepts and results on controllability and observability, can skip Sections 6.2–6.5.

6.2 CONTROLLABILITY: DEFINITIONS AND BASIC RESULTS

In this section, we introduce the basic concepts and some algebraic criteria of controllability.

6.2.1 Controllability of a Continuous-Time System

Definition 6.2.1. *The system:*

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t) + Du(t) \end{aligned} \tag{6.2.1}$$

is said to be **controllable**, *if starting from any initial state $x(0)$, the system can be driven to any final state $x_1 = x(t_1)$ in some finite time t_1, choosing the input variable $u(t), 0 \leq t \leq t_1$ appropriately.*

Remark

- The controllability of the system (6.2.1) is often referred to as the controllability of the pair (A, B), the reason for which will be clear in the following theorem.

Theorem 6.2.1. *Criteria for Continuous-Time Controllability. Let $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m} (m \leq n)$.*

The following are equivalent:

(i) *The system* (6.2.1) *is controllable.*

(ii) *The $n \times nm$ matrix*

$$C_{\mathrm{M}} = (B, AB, A^2B, \ldots, A^{n-1}B)$$

has full rank n.

(iii) *The matrix*

$$W_{\mathrm{C}} = \int_0^{t_1} e^{At} B B^{\mathrm{T}} e^{A^{\mathrm{T}} t}\, dt$$

is nonsingular for any $t_1 > 0$.

(iv) *If (λ, x) is an eigenpair of A^{T}, that is, $x^{\mathrm{T}}A = \lambda x^{\mathrm{T}}$, then $x^{\mathrm{T}}B \neq 0$*

(v) *Rank $(A - \lambda I, B) = n$ for every eigenvalue λ of A.*

(vi) *The eigenvalues of $A - BK$ can be arbitrarily assigned (assuming that the complex eigenvalues occur in conjugate pairs) by a suitable choice of K.*

Proof. Without loss of generality, we can assume that $t_0 = 0$. Let $x(0) = x_0$.

(i) ⇒ (ii). Suppose that the rank of C_{M} is not n. From Chapter 5, we know that

$$x(t_1) = e^{At_1}x_0 + \int_0^{t_1} e^{A(t_1-t)} Bu(t)\, dt. \tag{6.2.2}$$

That is,

$$\begin{aligned} x(t_1) - e^{At_1}x_0 &= \int_0^{t_1} \left\{ I + A(t_1 - t) + \frac{A^2}{2!}(t_1 - t)^2 + \cdots \right\} Bu(t)\, dt \\ &= B\int_0^{t_1} u(t)\, dt + AB\int_0^{t_1} (t_1 - t)u(t)\, dt + A^2B\int_0^{t_1} \frac{(t_1 - t)^2}{2!} u(t)\, dt + \cdots \end{aligned}$$

From the Cayley–Hamilton Theorem (see **Chapter 1**), it then follows that the vector $x(t_1)$ is a linear combination of the columns of $B, AB, A^2B, \ldots, A^{n-1}B$.

Since C_{M} does not have rank n, these columns vectors cannot form a basis of the state-space and therefore for some t_1, $x(t_1) = x_1$ cannot be attained, implying that (6.2.1) is not controllable.

(ii) ⇒ (iii). Suppose that the matrix C_{M} has rank n, but the matrix:

$$W_{\mathrm{C}} = \int_0^{t_1} e^{At} B B^{\mathrm{T}} e^{A^{\mathrm{T}} t}\, dt \tag{6.2.3}$$

is singular.

Let v be a nonzero vector such that $W_{\mathrm{C}}v = 0$. Then, $v^{\mathrm{T}} W_{\mathrm{C}}\, v = 0$. That is, $\int_0^{t_1} v^{\mathrm{T}} e^{At} B B^{\mathrm{T}} e^{A^{\mathrm{T}} t} v\, dt = 0$. The integrand is always nonnegative, since it is of the form $c^{\mathrm{T}}(t)c(t)$, where $c(t) = B^{\mathrm{T}} e^{A^{\mathrm{T}} t} v$. Thus, for the above integral to be equal to zero, we must have:

$$v^{\mathrm{T}} e^{At} B = 0, \quad \text{for } 0 \le t \le t_1.$$

From this we obtain (evaluating the successive derivatives with respect to t, at $t = 0$):

$$v^{\mathrm{T}} A^i B = 0, \quad i = 1, 2, \ldots, n-1.$$

That is, v is orthogonal to all columns of the matrix C_{M}. Since it has been assumed that the matrix C_{M} has rank n, this means that $v = 0$, which is a contradiction.

(iii) ⇒ (i). We show that $x(t_1) = x_1$. Let us now choose the vector $u(t)$ as

$$u(t) = B^{\mathrm{T}} e^{A^{\mathrm{T}}(t_1 - t)} W_{\mathrm{C}}^{-1}(-e^{At_1} x_0 + x_1).$$

Then from (6.2.2), it is easy to see that $x(t_1) = x_1$. This implies that the system (6.2.1) is controllable.

(ii) ⇒ (iv). Let x be an eigenvector of A^{T} corresponding to an eigenvalue λ; that is, $x^{\mathrm{T}} A = \lambda x^{\mathrm{T}}$. Suppose that $x^{\mathrm{T}} B = 0$. Then,

$$x^{\mathrm{T}} C_{\mathrm{M}} = (x^{\mathrm{T}} B, \lambda x^{\mathrm{T}} B, \lambda^2 x^{\mathrm{T}} B, \ldots, \lambda^{n-1} x^{\mathrm{T}} B) = 0.$$

Since the matrix C_{M} has full rank, $x = 0$, which is a contradiction.

(iv) ⇒ (ii). Assume that none of the eigenvectors of A^{T} is orthogonal to the columns of B, but rank $(C_{\mathrm{M}}) = k < n$.

We will see later in this chapter (Theorem 6.4.1) that, in this case, there exists a nonsingular matrix T such that

$$\bar{A} = TAT^{-1} = \begin{pmatrix} \bar{A}_{11} & \bar{A}_{12} \\ 0 & \bar{A}_{22} \end{pmatrix}, \qquad \bar{B} = TB = \begin{pmatrix} \bar{B}_1 \\ 0 \end{pmatrix}, \tag{6.2.4}$$

where $\bar{A}_{22}$ is of order $(n-k)$, and $k = \mathrm{rank}(C_{\mathrm{M}})$.

Let v_2 be an eigenvector of $(\bar{A}_{22})^{\mathrm{T}}$ corresponding to an eigenvalue λ. Then,

$$(\bar{A})^{\mathrm{T}} \begin{pmatrix} 0 \\ v_2 \end{pmatrix} = \begin{pmatrix} \bar{A}_{11}^{\mathrm{T}} & 0 \\ \bar{A}_{12}^{\mathrm{T}} & \bar{A}_{22}^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} 0 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ \bar{A}_{22}^{\mathrm{T}} v_2 \end{pmatrix} = \lambda \begin{pmatrix} 0 \\ v_2 \end{pmatrix}.$$

Furthermore,

$$(0, v_2^{\mathrm{T}}) \bar{B} = (0, v_2^{\mathrm{T}}) \begin{pmatrix} \bar{B}_1 \\ 0 \end{pmatrix} = 0.$$

That is, there is an eigenvector, namely $\binom{0}{v_2}$ of $(\bar{A})^{\mathrm{T}}$ such that it is orthogonal to the columns of $\bar{B}$. This means that the pair $(\bar{A}, \bar{B})$ is not controllable.

This is a contradiction because a similarity transformation does not change controllability.

(ii) $\Rightarrow$ (v). Rank$(\lambda I - A, B) < n$ if and only if there exists a nonzero vector v such that $v^{\mathrm{T}}(\lambda I - A, B) = 0$.

This equation is equivalent to:

$$A^{\mathrm{T}} v = \lambda v \quad \text{and} \quad v^{\mathrm{T}} B = 0.$$

This means that v is an eigenvector of A^{T} corresponding to the eigenvalue λ and it is orthogonal to the columns of B. The system (A, B) is, therefore, not controllable by (iv).

(v) $\Rightarrow$ (ii). If (v) were not true, then from (iv), we would have had

$$x^{\mathrm{T}}(B, AB, \ldots, A^{n-1}B) = 0,$$

meaning that rank(C_{M}) is less than n.

(vi) $\Rightarrow$ (i). Suppose that (vi) holds, but not (i). Then the system can be decomposed into (6.2.4) such that a subsystem corresponding to $\bar{A}_{22}$ is uncontrollable, whose eigenvalues, therefore, cannot be changed by the control. This contradicts (vi).

(i) $\rightarrow$ (vi). The proof of this part will be given in Chapter 10 (**Theorem 10.4.1**). It will be shown there that if (A, B) is controllable, then a matrix K can be constructed such that the eigenvalues of the matrix $(A - BK)$ are in desired locations. ■

Definition 6.2.2. *The matrix*

$$C_{\mathrm{M}} = (B, AB, A^2B, \ldots, A^{n-1}B) \tag{6.2.5}$$

is called the **controllability matrix**.

Remark

- The eigenvector criterion (iv) and the eigenvalue criterion (v) are popularly known as the **Popov–Belevitch–Hautus (PBH)** criteria of controllability (see Hautus 1969). For a historical perspective of this title, see Kailath (1980, p. 135).

Component controllability. The controllability, as defined in the Definition 6.2.1, is often referred to as the **complete controllability** implying that all the states are controllable.

However, if only one input, say $u_j(t)$, from $u(t) = (u_1(t), \ldots, u_m(t))^{\mathrm{T}}$ is used, then the rank of the corresponding $n \times n$ controllability matrix $C_{\mathrm{M}}^j = (b_j, Ab_j, \ldots, A^{n-1}b_j)$, where b_j is the jth column of B, determines the number of states that are controllable using the input $u_j(t)$. This is illustrated in the following example.

Consider Example 5.2.6 on the motions of an orbiting satellite with $d_0 = 1$.

It is easy to see that C_M has rank 4, so that all states are controllable using both inputs. However, if only the first input $u_1(t)$ is used, then

$$C_M^1 = (b_1, Ab_1, A^2b_1, A^3b_1) = \begin{pmatrix} 0 & 1 & 0 & -\omega^2 \\ 1 & 0 & -\omega^2 & 0 \\ 0 & 0 & -2\omega & 0 \\ 0 & -2\omega & 0 & 2\omega^3 \end{pmatrix},$$

which is singular.

Thus, one of the states is not controllable by using only the radial force $u_1(t)$.

However, it can be easily verified that all the states would be controllable if the tangential force $u_2(t)$ were used instead of $u_1(t)$.

Controllable and Uncontrollable Modes

From the eigenvalue and eigenvector criteria above, it is clear that the controllability and uncontrollability of the pair (A, B) are tied with the eigenvalues and eigenvectors of the system matrix A.

Definition 6.2.3. *A mode of the system* (6.2.1) *or, equivalently, an eigenvalue λ of A is controllable if the associated left eigenvector (i.e., the eigenvector of A^T associated with λ) is not orthogonal to the columns of B. Otherwise, the mode is uncontrollable.*

6.2.2 Controllability of a Discrete-Time System

Definition 6.2.4. *The discrete-time system*

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k, \\ y_k &= Cx_k + Du_k \end{aligned} \tag{6.2.6}$$

is said to be **controllable** *if for any initial state x_0 and any final state x_1, there exists a finite sequence of inputs $\{u_0, u_1, \cdots, u_{N-1}\}$ that transfers x_0 to x_1; that is, $x(N) = x_1$.*

In particular, if $x_0 = 0$ and the system (6.2.6) *is controllable, then it is called* ***reachable*** *(see Chen* 1984). *It is also known as controllability from the origin.*

Note: To avoid any confusion, we will assume (without any loss of generality) that $x_0 = 0$. So, in our case, the notion of controllability and reachability are equivalent.

Most of the criteria on controllability in the continuous-time case also hold in the discrete-time case. Here we will state and prove only one criterion analogous to (ii) of Theorem 6.2.1.

Theorem 6.2.2. *The discrete-time system* (6.2.6) *or the pair* (A, B) *is controllable if and only if the rank of the controllability matrix*

$$C_{\mathrm{M}} = (B, AB, \ldots, A^{n-1}B)$$

is n.

Proof. From Theorem 5.4.1, we know that the general solution of the discrete-time systems is

$$x_N = A^{N-1}Bu_0 + A^{N-2}Bu_1 + \cdots + Bu_{N-1}.$$

Thus, x_N can be expressed as a linear combination of $A^{k-1}B$, $k = N, \ldots, 1$.

So, it is possible to choose u_0 through u_{N-1} for arbitrary x_N if and only if the sequence $\{A^i B\}$ has a finite number of columns that span $\mathbb{R}^n$; and this is possible, if and only if the controllability matrix C_{M} has rank n. ■

6.3 OBSERVABILITY: DEFINITIONS AND BASIC RESULTS

In this section we state definitions and basic results of observability. The results will not be proved here because they can be easily proved by duality of the results on controllability proved in the previous section.

6.3.1 Observability of a Continuous-Time System

The concept of observability is dual to the concept of controllability.

Definition 6.3.1. *The continuous-time system* (6.2.1) *is said to be* **observable** *if there exists* $t_1 > 0$ *such that the initial state* $x(0)$ *can be uniquely determined from the knowledge of* $u(t)$ *and* $y(t)$ *for all* t, $0 \leq t \leq t_1$.

Remark

- The observability of the system (6.2.1) is often referred to as the observability of the pair (A, C).

Analogous to the case of controllability, we state the following criteria of observability.

Theorem 6.3.1. Criteria for Continuous-Time Observability. *The following are equivalent:*

(i) *The system* (6.2.1) *is observable.*

(ii) *The observability matrix*

$$O_M = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix}$$

has full rank n.

(iii) *The matrix*

$$W_O = \int_0^{t_1} e^{A^T\tau} C^T C e^{A\tau} d\tau$$

is nonsingular for any $t_1 > 0$.

(iv) *The matrix*

$$\begin{bmatrix} \lambda I - A \\ C \end{bmatrix}$$

has rank n for every eigenvalue λ *of A.*

(v) *None of the eigenvectors of A is orthogonal to the rows of C, that is, if* (λ, y) *is an eigenpair of A, then* $Cy \neq 0$.

(vi) *There exists a matrix L such that the eigenvalues of* $A + LC$ *can be assigned arbitrarily, provided that the complex eigenvalues occur in conjugate pairs.*

We only prove (iii) $\Longleftrightarrow$ (i) *and leave the others as an exercise* **(Exercise 6.6)**.

Theorem 6.3.2. *The pair* (A, C) *is observable if and only if the matrix* W_O *is nonsingular for any* $t_1 > 0$.

Proof. First suppose that the matrix W_O is nonsingular. Since $y(t)$ and $u(t)$ are known, we can assume, without any loss of generality, that $u(t) = 0$ for every t. Thus,

$$y(t) = Ce^{At}x(0).$$

This gives

$$W_O x(0) = \int_0^{t_1} e^{A^T\tau} C^T y(\tau)\, d\tau.$$

Thus, $x(0)$ is uniquely determined and is given by $x(0) = W_O^{-1} \int_0^{t_1} e^{A^T\tau} C^T y(\tau) d\tau$.

Conversely, if W_O is singular, then there exists a nonzero vector z such that $W_O z = 0$, which in turn implies that $Ce^{A\tau}z = 0$. So, $y(\tau) = Ce^{A\tau}(x(0) + z) = Ce^{A\tau}x(0)$.

Thus, $x(0)$ cannot be determined uniquely, implying that (A, C) is not observable. ∎

Component observability. As in the case of controllability, we can also speak of component observability when all the states are not observable by certain output. The rank of the observability matrix

$$C_{\mathrm{M}}^{j} = \begin{pmatrix} c_j \\ c_j A \\ \vdots \\ c_j A^{n-1} \end{pmatrix},$$

where c_j, the jth row of the output matrix C, determines the number of states that are observable by the output $y_j(t)$.

6.3.2 Observability of a Discrete-Time System

Definition 6.3.2. *The discrete-time system* (6.2.6) *is said to be observable if there exists an index N such that the initial state x_0 can be completely determined from the knowledge of inputs $u_0, u_1, \ldots, u_{N-1}$, and the outputs $y_0, y_1, \ldots, y_N$.*

The criteria of observability in the discrete-time case are the same as in the continuous-time case, and therefore, will not be repeated here.

6.4 DECOMPOSITIONS OF UNCONTROLLABLE AND UNOBSERVABLE SYSTEMS

Suppose that the pair (A, B) is not controllable. Let the rank of the controllability matrix be $k < n$. Then the following theorem shows that the system can be decomposed into controllable and uncontrollable parts.

Theorem 6.4.1. *Decomposition of Uncontrollable System. If the controllability matrix C_{M} has rank k, then there exists a nonsingular matrix T such that*

$$\bar{A} = TAT^{-1} = \begin{pmatrix} \bar{A}_{11} & \bar{A}_{12} \\ 0 & \bar{A}_{22} \end{pmatrix}, \qquad \bar{B} = TB = \begin{pmatrix} \bar{B}_1 \\ 0 \end{pmatrix}, \tag{6.4.1}$$

where $\bar{A}_{11}$, $\bar{A}_{12}$, and $\bar{A}_{22}$ are, respectively, of order $k \times k$, $k \times (n-k)$, and $(n-k) \times (n-k)$, and $\bar{B}_1$ has k rows. Furthermore, $(\bar{A}_{11}, \bar{B}_1)$ is controllable.

Proof. Let $v_1, \ldots, v_k$ be the independent columns of the controllability matrix C_{M}. We can always choose a set of $n-k$ vectors $v_{k+1}, \ldots, v_n$ so that the vectors $(v_1, v_2, \ldots, v_k, v_{k+1}, \ldots, v_n)$ form a basis of $\mathbb{R}^n$.

Then it is easy to see that the matrix $T^{-1} = (v_1, \ldots, v_n)$ is such that TAT^{-1} and TB will have the above forms.

To prove that $(\bar{A}_{11}, \bar{B}_1)$ is controllable, we note that the controllability matrix of the pair $(\bar{A}, \bar{B})$ is

$$\begin{pmatrix} \bar{B}_1 & \bar{A}_{11}\bar{B}_1 & \cdots & (\bar{A}_{11})^{k-1}\bar{B}_1 & \cdots & (\bar{A}_{11})^{n-1}\bar{B}_1 \\ 0 & 0 & & 0 & \cdots & 0 \end{pmatrix}.$$

Since, for each $j \geq k$, $(\bar{A}_{11})^j$ is a linear combination of $(\bar{A}_{11})^i$, $i = 0, 1, \ldots, (k-1)$, by the Cayley–Hamilton Theorem (see Chapter 1), we then have $\operatorname{rank}(\bar{B}_1, \bar{A}_{11}\bar{B}_1, \ldots, \bar{A}_{11}^{k-1}\bar{B}_1) = k$, proving that $(\bar{A}_{11}, \bar{B}_1)$ is controllable. ■

Note: If we define $\bar{x} = Tx$, then the state vector $\bar{x}$ corresponding to the system defined by $\bar{A}$ and $\bar{B}$ is given by $\bar{x} = \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \end{pmatrix}$.

Remark (Choosing T Orthogonal)

- Note that T in Theorem 6.4.1 can be chosen to be orthogonal by finding the QR factorization of the controllability matrix C_{M}. Thus, if $C_{\mathrm{M}} = QR$, then $T = Q^{\mathrm{T}}$.

Using duality, we have the following decomposition for an unobservable pair. The proof is left as an exercise **(Exercise 6.7)**.

Theorem 6.4.2. *Decomposition of Unobservable System. If the observability matrix O_{M} has rank $k' < n$, then there exists a nonsingular matrix T such that*

$$\bar{A} = TAT^{-1} = \begin{pmatrix} \bar{A}_{11} & \bar{A}_{12} \\ 0 & \bar{A}_{22} \end{pmatrix}, \qquad \bar{C} = CT^{-1} = (0, \bar{C}_1) \qquad (6.4.2)$$

with $(\bar{A}_{11}, \bar{C}_1)$ observable and $\bar{A}_{11}$ is of order k'.

The Kalman Decomposition

Combining Theorems 6.4.1 and 6.4.2, we obtain (after some reshuffling of the coordinates) the following decomposition, known as the **Kalman Canonical Decomposition**. The proof of the theorem is left as an exercise **(Exercise 6.8)**, and can also be found in any standard text on linear systems theory.

Theorem 6.4.3. *The Kalman Canonical Decomposition Theorem. Given the system* (6.2.1) *there exists a nonsingular coordinate transformation $\bar{x} = Tx$*

such that

$$\begin{pmatrix} \dot{\bar{x}}_{c\bar{o}} \\ \dot{\bar{x}}_{co} \\ \dot{\bar{x}}_{\bar{c}\bar{o}} \\ \dot{\bar{x}}_{\bar{c}o} \end{pmatrix} = \begin{pmatrix} \bar{A}_{c\bar{o}} & \bar{A}_{12} & \bar{A}_{13} & \bar{A}_{14} \\ 0 & \bar{A}_{co} & 0 & \bar{A}_{24} \\ 0 & 0 & \bar{A}_{\bar{c}\bar{o}} & \bar{A}_{34} \\ 0 & 0 & 0 & \bar{A}_{\bar{c}o} \end{pmatrix} \begin{pmatrix} \bar{x}_{c\bar{o}} \\ \bar{x}_{co} \\ \bar{x}_{\bar{c}\bar{o}} \\ \bar{x}_{\bar{c}o} \end{pmatrix} + \begin{pmatrix} \bar{B}_{c\bar{o}} \\ \bar{B}_{co} \\ 0 \\ 0 \end{pmatrix}, \tag{6.4.3}$$

$$y = (0, \bar{C}_{co}, 0, \bar{C}_{\bar{c}o}) \begin{pmatrix} \bar{x}_{c\bar{o}} \\ \bar{x}_{co} \\ \bar{x}_{\bar{c}\bar{o}} \\ \bar{x}_{\bar{c}o} \end{pmatrix} + Du,$$

$\bar{x}_{c\bar{o}} \equiv$ *states which are controllable but not observable.*
$\bar{x}_{co} \equiv$ *states which are both controllable and observable*
$\bar{x}_{\bar{c}\bar{o}} \equiv$ *states which are both uncontrollable and unobservable*
$\bar{x}_{\bar{c}o} \equiv$ *states which are uncontrollable but observable.*
Moreover, the transfer function matrix from u to y is given by

$$G(s) = \bar{C}_{co}(sI - \bar{A}_{co})^{-1}\bar{B}_{co} + D.$$

Remark

- It is interesting to observe that the uncontrollable and/or unobservable parts of the system do not appear in the description of the transfer function matrix.

6.5 CONTROLLER- AND OBSERVER-CANONICAL FORMS

An important property of a linear system is that **controllability and observability remain invariant under certain transformations**. We will state the result for controllability without proof. A similar result, of course, holds for observability. Proof is left as an **Exercise (6.9)**.

Theorem 6.5.1. *Let T be a nonsingular matrix such that $TAT^{-1} = \tilde{A}$, and $TB = \tilde{B}$, then (A, B) is controllable if and only if $(\tilde{A}, \tilde{B})$ is controllable.*

The question naturally arises if the matrix T can be chosen so that $\tilde{A}$ and $\tilde{B}$ will have simple forms, from where conclusions about controllability or observability can be easily drawn. Two such forms, **controller-canonical form** (or the **controller-companion form**) and the **Jordan Canonical Form** (JCF) are well known in control text books (Luenberger 1979; Kailath 1980; Chen 1984; Szidarovszky and Bahill 1991 etc.). **Unfortunately, neither of these two forms, in general, can be obtained in a numerically stable way, because, T, not being an orthogonal matrix in general, can be highly ill-conditioned.**

The controller- and observer-canonical forms are, *however, very valuable theoretical tools in establishing many theoretical results in control and systems theory* (e.g., the proof of the eigenvalue assignment (EVA) theorem by state feedback (**Theorem 10.4.1 in Chapter 10**)). We just state these two forms here for later uses as theoretical tools. First consider the single-input case.

Let (A, b) be controllable and let C_{M} be the controllability matrix.

Let s_n be the last row of C_{M}^{-1}. Then the matrix T defined by

$$T = \begin{pmatrix} s_n \\ s_n A \\ \vdots \\ s_n A^{n-1} \end{pmatrix} \tag{6.5.1}$$

is such that

$$\tilde{A} = TAT^{-1} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & & 1 \\ -a_1 & -a_2 & -a_3 & \cdots & -a_n \end{pmatrix}, \qquad \tilde{b} = Tb = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \tag{6.5.2}$$

Similarly, it is possible to show that if (A, b) is controllable, then there exists a nonsingular matrix P such that

$$PAP^{-1} = \begin{pmatrix} -a_n & -a_{n-1} & \cdots & -a_2 & -a_1 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}, \qquad Pb = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{6.5.3}$$

The forms (6.5.1) and (6.5.3) are, respectively, known as **lower** and **upper companion (or controller) canonical forms.** By duality, the **observer-canonical forms** (in lower and upper companion forms) can be defined. Thus, the pair $(\tilde{A}, \tilde{c})$ given by

$$\tilde{A} = \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_1 \\ 1 & 0 & \cdots & 0 & -a_2 \\ 0 & 1 & \cdots & 0 & -a_3 \\ \vdots & & & & \vdots \\ 0 & 0 & \cdots & 1 & -a_n \end{pmatrix}, \qquad \tilde{c} = (0, 0, \ldots, 0, 1).$$

is an **upper observer-canonical form**.

MATCONTROL note: The MATCONTROL function **cntrlc** can be used to find a controller-canonical form.

The Luenberger Canonical Form

In the multi-input case, the controllable pair (A, B) can be reduced to the pair $(\tilde{A}, \tilde{B})$, where $\tilde{A} = TAT^{-1}$ is a block matrix, *in which each diagonal block matrix is a companion matrix* of the form given in (6.5.2), and $\tilde{B}$ is also a block matrix with each block as a companion matrix of the form (6.5.2) having nonzero entries only on the last row.

The number of diagonal blocks of $\tilde{A}$ is equal to the rank of B. Such a form is known as the **Luenberger controller-canonical form**. Similarly, by duality, the **Luenberger observer-canonical form** can be written down.

Numerical instability: In general, like a controller-companion form, the Luenberger canonical form also cannot be obtained in a numerically stable way.

6.6 NUMERICAL DIFFICULTIES WITH THEORETICAL CRITERIA OF CONTROLLABILITY AND OBSERVABILITY

Each of the algebraic criterion of controllability (observability) described in Section 6.2 (Section 6.3) suggests a test of controllability (observability). Unfortunately, most of them do not lead to numerically viable tests as the following discussions show. First, let's look into the controllability criterion (ii) of Theorem 6.2.1.

This criterion requires successive matrix multiplications and determination of the rank of an $n \times nm$ matrix.

It is well known that matrix multiplications involving nonorthogonal matrices can lead to a severe loss of accuracy (see Chapter 3). The process may transform the problem to a more sensitive one. To illustrate this, consider the following illuminating example from Paige (1981).

Example 6.6.1

$$A = \begin{pmatrix} 1 & & & \\ & 2^{-1} & & \\ & & \ddots & \\ & & & 2^{-9} \end{pmatrix}_{10\times 10}, \qquad B = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \cdots \\ 1 \end{pmatrix}.$$

The pair (A, B) is clearly controllable. The controllability matrix $(B, AB, \ldots, A^9B)$ can be computed easily and stored accurately. Note that the (i, j)-th entry of this matrix is $2^{(-i+1)(j-1)}$. This matrix has three smallest singular values 0.613×10^{-12}, 0.364×10^{-9}, and 0.712×10^{-7}. Thus, on a computer with machine precision no smaller than 10^{-12}, one will conclude that the *numerical rank of this matrix* is less than 10, indicating that the system is uncontrollable. (Recall that matrix A is said to have a *numerical rank* r if the computed singular values $\tilde{\sigma}_i, i = 1, \ldots, n$ satisfy $\tilde{\sigma}_1 \geq \tilde{\sigma}_2 \geq \cdots \geq \tilde{\sigma}_r \geq \delta \geq \tilde{\sigma}_{r+1} \geq \cdots \geq \tilde{\sigma}_n$, where δ is a tolerance) **(Section 3.9.4)**.

Note that determining the rank of a matrix using the singular values is the most effective way from a numerical viewpoint.

The criteria (iv)–(vi) in Theorem 6.2.1 are based on eigenvalues and eigenvectors computations. We know that the eigenvalues and eigenvectors of certain matrices can be very ill-conditioned and that the ill-conditioning of the computed eigenvalues will again lead to inaccuracies in computations. For example, by criteria (vi) of controllability in Theorem 6.2.1, it is possible, when (A, B) is controllable, to find a matrix K such that $A + BK$ and A have disjoint spectra. Computationally, it is indeed a difficult task to decide if two matrices have a common eigenvalue if that eigenvalue is ill-conditioned. This can be seen from the following discussion.

Let λ and δ be the eigenvalues of A and $A + BK$, respectively. We know that a computed eigenvalue $\tilde{\lambda}$ of A is an eigenvalue of the perturbed matrix $\tilde{A} = A + \Delta A$, where $\|\Delta A\|_2 \leq \mu \|A\|_2$. Similarly, a computed eigenvalue $\tilde{\delta}$ of $A + BK$ is an eigenvalue of $A + BK + \Delta A'$, where $\|\Delta A'\|_2 \leq \mu \|A + BK\|_2$. Thus, even if $\lambda = \delta$, $\tilde{\lambda}$ can be very different from λ and $\tilde{\delta}$ very different from δ, implying that $\tilde{\lambda}$ and $\tilde{\delta}$ are different.

Example 6.6.2. Consider the following example due to Paige (1981) where the matrices A and B are taken as

$$A = Q^{\mathrm{T}} \hat{A} Q, \qquad B = Q^{\mathrm{T}} \hat{B}$$

Here $\hat{A}$ is the well-known 20×20 Wilkinson bidiagonal matrix

$$\hat{A} = \begin{pmatrix} 20 & 20 & & 0 & \\ & 19 & 20 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 20 \\ 0 & & & & 1 \end{pmatrix},$$

$$\hat{B} = (1, 1, \ldots, 1, 0)^{\mathrm{T}},$$

and Q is the Q-matrix of the QR factorization of a randomly generated 20×20 arbitrary matrix whose entries are uniform random numbers on $(-1, 1)$.

Clearly, the pair $(\hat{A}, \hat{B})$, and therefore, the pair (A, B), are uncontrollable. (Note that controllability or uncontrollability is preserved by nonsingular transformations (Theorem 6.5.1)).

Now taking K as a 1×20 matrix with entries as random numbers uniformly distributed on $(-1, 1)$, the eigenvalues $\tilde{\lambda}_i$ of A and μ_i of $A + BK$ were computed and tabulated. They are displayed in the following table.

In this table, $\rho(B, A - \tilde{\lambda}_i I)$ denotes the ratio of the smallest to the largest singular value of the matrix $(B, A - \tilde{\lambda}_i I)$.

Eigenvalues $\tilde{\lambda}_i(A)$	Eigenvalues $\mu_i(A+BK)$	$\rho(B, A-\tilde{\lambda}_i I)$
$-0.32985 \pm j1.06242$	$0.99999 \pm j0$	0.002
$0.9219 \pm j3.13716$	$-8.95872 \pm j3.73260$	0.004
$3.00339 \pm j3.13716$	$-5.11682 \pm j9.54329$	0.007
$5.40114 \pm j6.17864$	$-0.75203 \pm j14.148167$	0.012
$8.43769 \pm j7.24713$	$5.77659 \pm j15.58436$	0.018
$11.82747 \pm j7.47463$	$11.42828 \pm j14.28694$	0.026
$15.10917 \pm j6.90721$	$13.30227 \pm j12.90197$	0.032
$18.06886 \pm j5.66313$	$18.59961 \pm j14.34739$	0.040
$20.49720 \pm j3.81950$	$23.94877 \pm j11.80677$	0.052
$22.06287 \pm j1.38948$	$28.45618 \pm j8.45907$	0.064

The table shows that $\tilde{\lambda}_i$ are almost unrelated to μ_i. One will, then, erroneously conclude that the pair (A, B) is controllable.

The underlying problem, of course, is the ill-conditioning of the eigenvalues of $\hat{A}$. Note that, because of ill-conditioning, the computed eigenvalues of A are different from those of $\hat{A}$, which, in theory, should have been the same because A and $\hat{A}$ are similar.

The entries of the third column of the table can be used to illustrate the difficulty with the eigenvalue criterion (Criterion (v) of Theorem 6.2.1).

Since the pair (A, B) is uncontrollable, by the eigenvalue criterion of controllability, $\text{rank}(B, A-\tilde{\lambda}_i I)$, for some $\tilde{\lambda}_i$, should be less than n; consequently, one of the entries of the third column should be identically zero. But this is not the case; **only there is an indication that some "modes" are less controllable than the others.**

To confirm the fact that ill-conditioning of the eigenvalues of $\hat{A}$ is indeed the cause of such failure, $\text{rank}(B, A-I)$, which corresponds to the exact eigenvalue 1 of $\hat{A}$ was computed and seen to be

$$\text{rank}(B, A-I) = 5 \times 10^{-8}.$$

Thus, this test would have done well if the exact eigenvalues of $\hat{A}$ were used in place of the computed eigenvalues of A, which are complex.

6.7 A NUMERICALLY EFFECTIVE TEST OF CONTROLLABILITY

A numerically effective test of controllability can be obtained through the reduction of the pair (A, B) to a *block Hessenberg form* using **orthogonal similarity transformation**. The process constructs an orthogonal matrix P such that

$$\begin{aligned} PAP^{\mathrm{T}} &= H, \text{ a block upper Hessenberg matrix} \\ PB &= \tilde{B} = \begin{pmatrix} B_1 \\ 0 \end{pmatrix}. \end{aligned} \tag{6.7.1}$$

The form (6.7.1) is called the **controller-Hessenberg** form of (A, B), and the pair $(H, \bar{B})$ is called the controller-Hessenberg pair of (A, B) (see Rosenbrock 1970). The reduction to this form can be done using Householder's or Givens' method. We describe the reduction here using Householder's transformations. The algorithmic procedure appears in Boley (1981), Van Dooren and Verhaegen (1985), and Paige (1981), Patel (1981), Miminis (1981) etc. The algorithm is usually known as the **staircase algorithm**.

Algorithm 6.7.1. *Staircase Algorithm. Let A be $n \times n$ and B be $n \times m$ $(m \leq n)$.*

Step 0. *Triangularize the matrix B using the QR factorization with column pivoting (Golub and Van Loan, 1996, pp. 248–250), that is, find an orthogonal matrix P_1 and a permutation matrix E_1 such that*

$$P_1 B E_1 = \begin{pmatrix} B_1 \\ 0 \end{pmatrix},$$

where B_1 is an $n_1 \times m$ upper triangular matrix and $n_1 = \text{rank}(B) = \text{rank}(B_1)$.

Step 1. *Update A and B, that is, compute*

$$P_1 A P_1^{\mathrm{T}} = H = \begin{pmatrix} H_{11}^{(1)} & H_{12}^{(1)} \\ H_{21}^{(1)} & H_{22}^{(1)} \end{pmatrix}, \qquad \tilde{B} = P_1 B = \begin{pmatrix} B_1 \\ 0 \end{pmatrix} E^{\mathrm{T}} \equiv \begin{pmatrix} \bar{B}_1 \\ 0 \end{pmatrix}.$$

where $H_{11}^{(1)}$ is $n_1 \times n_1$ and $H_{21}^{(1)}$ is $(n - n_1) \times n_1$, $n_1 \leq n$. If $H_{21}^{(1)} = 0$, stop.

Step 2. *Triangularize $H_{21}^{(1)}$ using the QR factorization with column pivoting, that is, find an orthogonal matrix $\hat{P}_2$ and a permutation matrix E_2 such that*

$$\hat{P}_2 H_{21}^{(1)} E_2 = \begin{pmatrix} H_{21}^{(*)} \\ 0 \end{pmatrix},$$

where $H_{21}^{(*)}$ is $n_2 \times n_1$, $n_2 = \mathrm{rank}(\mathrm{H}_{21}^{(1)}) = \mathrm{rank}(\mathrm{H}_{21}^{(*)})$, and $n_2 \leq n_1$.

If $n_1 + n_2 = n$, stop.

Form

$$P_2 = \text{diag}\,(I_{n_1}, \hat{P}_2) = \left(\begin{array}{c|c} I_{n_1} & 0 \\ \hline 0 & \hat{P}_2 \end{array}\right),$$

where I_{n_1} is a matrix consisting of the first n_1 rows and columns of the identity matrix.

Compute

$$H_2 = P_2 H_1 P_2^{\mathrm{T}} = \left(\begin{array}{ccc} H_{11}^{(1)} & H_{12}^{(2)} & H_{13}^{(2)} \\ \hline H_{21}^{(2)} & H_{22}^{(2)} & H_{23}^{(2)} \\ \hline 0 & H_{32}^{(2)} & H_{33}^{(2)} \end{array}\right),$$

where $H_{22}^{(2)}$ is $n_2 \times n_2$ and $H_{32}^{(2)}$ is $(n - n_1 - n_2) \times n_2$. Note that $H_{21}^{(2)} = H_{21}^{(*)} E_2^{\mathrm{T}}$.

Update $P \equiv P_2 P_1$.
If $H_{32}^{(2)} = 0$, *stop. (Note that* $H_{11}^{(1)}$ *does not change.)*

The matrix $\tilde{B}$ remains unchanged.

Step 3. *Triangularize* $H_{32}^{(2)}$ *to obtain its rank. That is, find* $\hat{P}_3$ *and* E_3 *such that*

$$\hat{P}_3 H_{32}^{(2)} E_3 = \begin{pmatrix} H_{32}^{(3)} \\ 0 \end{pmatrix}.$$

Let $n_3 = \text{rank}(H_{32}^{(2)}) = \text{rank}(H_{32}^{(3)})$; $n_3 \le n_2$.

If $n_1 + n_2 + n_3 = n$, *stop. Otherwise, compute* P_3, H_3, *and update* P *as above.* **(Note that $\tilde{B}$ remains unchanged.)**

Step 4. *Continue the process until for some integer* $k \le n$, *the algorithm produces*

$$H \equiv \begin{pmatrix} H_{11} & H_{12} & H_{13} & \cdots & H_{1k} \\ H_{21} & H_{22} & H_{23} & \cdots & H_{2k} \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & H_{k,k-1} & H_{kk} \end{pmatrix}, \qquad \tilde{B} \equiv \begin{pmatrix} B_1 \\ 0 \end{pmatrix}, \qquad (6.7.2)$$

where, either $H_{k,k-1}$ *has full rank* n_k, *signifying that the pair* (A, B) *is controllable, or* $H_{k,k-1}$ *is a zero matrix signifying that the pair* (A, B) *is uncontrollable.*

(Note that in the above expressions for H and $\tilde{B}$, the superscripts have been dropped, for convenience. *However,* H_{21} *stands for* $H_{21}^{(2)}$, H_{32} *stands for* $H_{32}^{(3)}$, *etc. that is,* $H_{k,k-1}$ *is established at step k).*

It is easy to see that

$$(\tilde{B}, H\tilde{B}, H^2\tilde{B}, \ldots, H^{k-1}\tilde{B}) = P(B, AB, \ldots, A^{k-1}B)$$
$$= \begin{pmatrix} B_1 & & \cdots & \cdots \\ & \boxed{H_{21}B_1} & \cdots & \cdots \\ & & \boxed{H_{32}H_{21}B_1} & \cdots \\ & & \ddots & \\ & & & \boxed{H_{k,k-1}\ldots H_{21}B_1} \end{pmatrix}$$

That is, it is block triangular matrix with $B_1, H_{21}B_1, \ldots, H_{k,k-1}, \ldots, H_{21}B_1$ *on the diagonal.*

This implies that the matrix $H_{k,k-1}$ *is of full rank if the system is controllable or is a zero matrix if the system is uncontrollable.*

Theorem 6.7.1. *Controller-Hessenberg Theorem. (i) Given the pair* (A, B), *the orthogonal matrix* P *constructed by the above procedure is such that*

$PAP^{\mathrm{T}} = H$ and $PB = \bar{B}$, where H and $\bar{B}$ are given by (6.7.2). *(ii) If pair (A, B) is controllable, then $H_{k,k-1}$ has full rank. If it is uncontrollable, then $H_{k,k-1} = 0$.*

Proof. The proof of Theorem 6.7.1 follows from the above construction. However, we will prove here part (ii) using (v) of Theorem 6.2.1.

Obviously, rank $(B, A - \lambda I) = n$ for all λ if and only if rank $(\bar{B}, H - \lambda I) = n$ for all λ.

Now,

$$(\tilde{B}, H - \lambda I) = \begin{pmatrix} B_1 & H_{11} - \lambda I_1 & \cdots & & \cdots & H_{1k} \\ 0 & H_{21} & & H_{22} - \lambda I_2 & \cdots & H_{2k} \\ \vdots & 0 & \ddots & & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & & \ddots \\ 0 & 0 & \cdots & 0 & H_{k,k-1} & H_{kk} - \lambda I_k \end{pmatrix}.$$

If the system is controllable, then the matrix $(\tilde{B}, H - \lambda I)$ must have full rank and thus, the matrix $H_{k,k-1}$ has full rank. On the other hand, if the system is not controllable, then the matrix $(\tilde{B}, H - \lambda I)$ cannot have full rank implying that $H_{k,k-1}$ must be a zero matrix. ■

Notes

1. The matrix $\bar{B}$ is not affected throughout the whole process.
2. At each step of computation, the rank of a matrix has to be determined. We have used QR factorization with column pivoting for this purpose. However, **the best way to do this is to use singular value decomposition (SVD) of that matrix.** (See Golub and Van Loan (1996) or Datta (1995)).
3. From the construction of the block Hessenberg pair $(H, \tilde{B})$, it follows that as soon as we encounter a zero block on the subdiagonal of H or if the matrix B_1 does not have full rank, we stop, concluding that (A, B) is not controllable.

Example 6.7.1. An Uncontrollable Pair

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Step 0. $P_1 = \begin{pmatrix} -0.5774 & -0.5774 & -0.5774 \\ 0.8165 & -0.4082 & -0.4082 \\ 0 & -0.7071 & 0.7071 \end{pmatrix}, \qquad E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$

$$P_1 B E_1 = \begin{pmatrix} -1.7321 & -1.7321 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Step 1. $H_1 = P_1 A P_1^{\mathrm{T}} = \begin{pmatrix} 2.3333 & 0.2357 & -0.4082 \\ -0.4714 & 0.1667 & -0.2887 \\ 0.8165 & -0.2887 & 0.5000 \end{pmatrix},$

$\tilde{B} = \begin{pmatrix} -1.7321 & -1.7321 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$

Step 2. $H_{21}^{(1)} = \begin{pmatrix} -0.4714 \\ 0.8165 \end{pmatrix}$

$$\hat{P}_2 = \begin{pmatrix} -0.5000 & 0.8660 \\ 0.8660 & 0.5000 \end{pmatrix}, \qquad E_2 = 1, \qquad P_2 = \mathrm{diag}(I_1, \hat{P}_2)$$

$$H_2 = P_2 H_1 P_2^{\mathrm{T}} = \begin{pmatrix} 2.3333 & -0.4714 & 0 \\ 0.9428 & 0.6667 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Since $H_{32}^{(2)} = 0$, we stop.

The controller-Hessenberg form is given by $(H \equiv H_2, \tilde{B})$

$$H = \begin{pmatrix} 2.3333 & -0.4714 & 0 \\ 0.9428 & 0.6667 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \tilde{B} = \begin{pmatrix} -1.7321 & -1.7321 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Clearly the pair (A, B) is not controllable.

Example 6.7.2. *A Controllable Pair*

$$A = \begin{pmatrix} 0.7665 & 0.1665 & 0.9047 & 0.4540 & 0.5007 \\ 0.4777 & 0.4865 & 0.5045 & 0.2661 & 0.3841 \\ 0.2378 & 0.8977 & 0.5163 & 0.0907 & 0.2771 \\ 0.2749 & 0.9092 & 0.3190 & 0.9478 & 0.9138 \\ 0.3593 & 0.0606 & 0.9866 & 0.0737 & 0.5297 \end{pmatrix},$$

$$B = \begin{pmatrix} 0.4644 & 0.8278 \\ 0.9410 & 0.1254 \\ 0.0501 & 0.0159 \\ 0.7615 & 0.6885 \\ 0.7702 & 0.8682 \end{pmatrix}.$$

Step 0.

$$P_1 = \begin{pmatrix} -0.3078 & -0.6236 & -0.0332 & -0.5047 & -0.5104 \\ 0.5907 & -0.7058 & -0.0263 & 0.1485 & 0.3610 \\ -0.0080 & -0.0451 & 0.9989 & -0.0047 & -0.0004 \\ -0.4561 & -0.2970 & -0.0132 & 0.8208 & -0.1728 \\ -0.5901 & -0.1510 & -0.0123 & -0.2225 & 0.7611 \end{pmatrix}, \quad E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

$$P_1 B E_1 = \begin{pmatrix} -1.5089 & -1.1241 \\ 0 & 0.8157 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

$n_1 = \text{rank}(B) = 2$

Step 1.

$$H_1 = P_1 A P_1^{\mathrm{T}} = \begin{pmatrix} 1.8549 & -0.3935 & -1.2228 & 0.1796 & -0.0198 \\ -0.3467 & 0.3934 & 0.5690 & 0.0593 & 0.0470 \\ -0.7857 & -0.4020 & 0.4421 & -0.3453 & -0.0848 \\ -0.5876 & -0.3573 & -0.4998 & 0.3311 & 0.2607 \\ 0.6325 & -0.0777 & 0.0837 & -0.1295 & 0.2253 \end{pmatrix}$$

$$= \begin{pmatrix} H_{11}^{(1)} & H_{12}^{(1)} \\ H_{21}^{(1)} & H_{22}^{(1)} \end{pmatrix}.$$

$$\tilde{B} = \begin{pmatrix} -1.5089 & -1.1241 \\ 0 & 0.8157 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Step 2.

$$\hat{P}_2 = \begin{pmatrix} -0.6731 & -0.5034 & 0.5418 \\ -0.3545 & -0.4233 & -0.8337 \\ 0.6490 & -0.7533 & 0.1064 \end{pmatrix}, \qquad E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

$$\hat{P}_2 H_{21}^{(1)} = \begin{pmatrix} H_{21}^{(2)} \\ 0 \end{pmatrix} = \begin{pmatrix} 1.1674 & 0.4084 \\ 0 & 0.3585 \\ 0 & 0 \end{pmatrix}.$$

$n_2 = \text{rank}(H_{21}^{(1)}) = \text{rank}(H_{21}^{(2)}) = 2.$

$$P_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -0.6731 & -0.5034 & 0.5418 \\ 0 & 0 & -0.3545 & -0.4233 & -0.8337 \\ 0 & 0 & 0.6490 & -0.7533 & 0.1064 \end{pmatrix},$$

$$P = \begin{pmatrix} -0.3078 & -0.6236 & -0.0332 & -0.5047 & -0.5104 \\ 0.5907 & -0.7058 & -0.0263 & 0.1485 & 0.3610 \\ -0.0847 & 0.0980 & -0.6724 & -0.5306 & 0.4996 \\ 0.6879 & 0.2676 & -0.3383 & -0.1602 & -0.5613 \\ 0.2756 & 0.1784 & 0.6570 & -0.6450 & 0.2109 \end{pmatrix}.$$

$$H_2 = P_2 H_1 P_2^{\mathrm{T}} = \begin{pmatrix} 1.8549 & -0.3935 & 0.7219 & 0.3740 & -0.9310 \\ -0.3467 & 0.3934 & -0.3874 & -0.2660 & 0.3297 \\ 1.1674 & 0.4084 & 0.0286 & -0.0378 & 0.0080 \\ 0 & 0.3585 & -0.1807 & 0.1907 & -0.1062 \\ 0 & 0 & -0.3304 & 0.1575 & 0.7792 \end{pmatrix}$$

$$= \begin{pmatrix} H_{11}^{(1)} & H_{12}^{(2)} & H_{13}^{(2)} \\ H_{21}^{(2)} & H_{22}^{(2)} & H_{23}^{(2)} \\ 0 & H_{32}^{(2)} & H_{33}^{(2)} \end{pmatrix}.$$

$$\bar{B} = \begin{pmatrix} -1.5089 & -1.1242 \\ 0 & 0.8157 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

$n_3 = \operatorname{rank}(H_{32}^{(2)}) = 1$.

Since $n_1 + n_2 + n_3 = 2 + 2 + 1 = 5$, we stop.

The controller-Hessenberg form (A, B) is given by $(H \equiv H_2, \bar{B})$.

The pair (A, B) is controllable, because $H_{21}^{(2)}$ and $H_{32}^{(2)}$ have full rank.

The next example (Example 6.7.3) shows the uses of non-identity premutation matrices in QR factorization with column pivoting.

Example 6.7.3.

$$A = \begin{pmatrix} 0.7665 & 0.1665 & 0.9047 & 0.4540 & 0.5007 \\ 0.4777 & 0.4865 & 0.5045 & 0.2661 & 0.3841 \\ 0.2378 & 0.8977 & 0.5163 & 0.0907 & 0.2771 \\ 0.2749 & 0.9092 & 0.3190 & 0.9478 & 0.9138 \\ 0.3593 & 0.0606 & 0.9866 & 0.0737 & 0.5297 \end{pmatrix},$$

$$B = \begin{pmatrix} 0.4644 & 0.8278 \\ 0.9410 & 0.1254 \\ 0.0501 & 1.0159 \\ 0.7615 & 0.6885 \\ 0.7702 & 0.8682 \end{pmatrix}.$$

Step 0.

$$P_1 = \begin{pmatrix} -0.4811 & -0.0729 & -0.5904 & -0.4001 & -0.5046 \\ 0.0213 & -0.7765 & 0.4917 & -0.3183 & -0.2312 \\ -0.6160 & 0.4529 & 0.6231 & -0.0783 & -0.1450 \\ -0.3829 & -0.3429 & -0.0646 & 0.8375 & -0.1739 \\ -0.4919 & -0.2627 & -0.1313 & -0.1764 & 0.8004 \end{pmatrix},$$

$$E_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Step 1.

$$\bar{B} = P_1 B = \begin{pmatrix} -1.0149 & -1.7207 \\ -1.1166 & -0.0000 \\ -0.0000 & 0.0000 \\ 0.0000 & 0.0000 \\ 0.0000 & 0.0000 \end{pmatrix},$$

$$n_1 = 2$$

$$H_1 = P_1 A P_1^{\mathrm{T}} = \begin{pmatrix} 2.1262 & 0.5717 & -0.6203 & 0.3595 & 0.1402 \\ 0.9516 & 0.2348 & -0.0160 & -0.0300 & -0.0472 \\ 0.2778 & -0.4685 & 0.3036 & -0.2619 & -0.0998 \\ 0.0115 & -0.8294 & 0.0243 & 0.3571 & 0.2642 \\ 0.3480 & 0.5377 & 0.0873 & -0.0673 & 0.2250 \end{pmatrix}.$$

Step 2.

$$\hat{P}_2 = \begin{pmatrix} -0.4283 & -0.7582 & 0.4916 \\ 0.6685 & 0.1002 & 0.7370 \\ -0.6080 & 0.6442 & 0.4640 \end{pmatrix},$$

$$E_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$n_2 = 2,$$

$$P_2 = \begin{pmatrix} 1.0000 & 0 & 0 & 0 & 0 \\ 0 & 1.0000 & 0 & 0 & 0 \\ 0 & 0 & -0.4283 & -0.7582 & 0.4916 \\ 0 & 0 & 0.6685 & 0.1002 & 0.7370 \\ 0 & 0 & -0.6080 & 0.6442 & 0.4640 \end{pmatrix},$$

$$P = P_2 P_1,$$

$$P = \begin{pmatrix} -0.4811 & -0.0729 & -0.5904 & -0.4001 & -0.5046 \\ 0.0213 & -0.7765 & 0.4917 & -0.3183 & -0.2312 \\ 0.3124 & -0.0631 & -0.2824 & -0.6881 & 0.5875 \\ -0.8127 & 0.0749 & 0.3133 & -0.0985 & 0.4755 \\ -0.1004 & -0.6182 & -0.4813 & 0.5053 & 0.3475 \end{pmatrix},$$

$$H_2 = P_2 H_1 P_2^{\mathrm{T}},$$

$$H_2 = \begin{pmatrix} 2.1262 & 0.5717 & 0.0619 & -0.2753 & 0.6738 \\ 0.9516 & 0.2348 & 0.0064 & -0.0485 & -0.0315 \\ 0.0434 & 1.0938 & 0.1675 & -0.1244 & -0.0811 \\ 0.4433 & 0.0000 & 0.0895 & 0.2539 & -0.2275 \\ -0.0000 & -0.0000 & -0.0517 & 0.1971 & 0.4644 \end{pmatrix},$$

$$\bar{B} = \begin{pmatrix} -1.0149 & -1.7207 \\ -1.1166 & -0.0000 \\ -0.0000 & 0.0000 \\ 0.0000 & 0.0000 \\ 0.0000 & 0.0000 \end{pmatrix}.$$

Flop-count: Testing controllability using the constructive proof of Theorem 6.7.1 requires roughly $6n^3 + 2n^2 m$ flops. The count includes the construction of the transforming matrix P (see Van Dooren and Verhaegen 1985).

Round-off error analysis and stability: The procedure is **numerically stable**. It can be shown that the computed matrices $\hat{H}$ and $\hat{B}$ are such that $\hat{H} = H + \Delta H$ and $\hat{B} = \bar{B} + \Delta B$, where $\|\Delta H\|_{\mathrm{F}} \le c\mu \|H\|_{\mathrm{F}}$ and $\|\Delta B\|_{\mathrm{F}} \le c\mu \|\|\bar{B}\|_{\mathrm{F}}$ for some small constant c. Thus, with the computed pair $(\hat{H}, \hat{B})$, we will compute the controllability of a system determined by the pair of matrices which are close to H and $\bar{B}$. Since the controllability of the pair $(H, \bar{B})$ is the same as that of the pair (A, B), **this can be considered as a backward stable method for finding the controllability of the pair** (A, B).

MATCONTROL note: Algorithm 6.7.1 has been implemented in MATCONTROL function **cntrlhs**.

The function **cntrlhst** gives block Hessenberg form with triangular subdiagonal blocks.

Controllability Index and Controller-Hessenberg Form

Let $B = (b_1, b_2, \ldots, b_m)$. Then the controllability matrix C_{M} can be written as

$$C_{\mathrm{M}} = (b_1, b_2, \ldots, b_m; Ab_1, Ab_2, \ldots, Ab_m; \ldots; A^{n-1}b_1, A^{n-1}b_2, \ldots, A^{n-1}b_m).$$

Suppose that the linearly independent columns of the matrix C_{M} have been obtained in order from left to right. Reorder these independent columns to obtain:

$$C'_{\mathrm{M}} = (b_1, Ab_1, \ldots, A^{\mu_1 - 1}b_1; b_2, Ab_2, \ldots, A^{\mu_2 - 1}b_2; \ldots; b_m, Ab_m, \ldots, A^{\mu_m - 1}b_m).$$

The integers $\mu_1, \ldots, \mu_m$ are called the **controllability indices** associated with $b_1, b_2, \ldots, b_m$, respectively if $\mu_1 \ge \cdots \ge \mu_m$. Note that μ_i is the number of independent columns associated with b_i.

Furthermore, $\mu = \max(\mu_1, \ldots, \mu_m)$ is called the **controllability index**. If $\mu_1 + \mu_2 + \cdots + \mu_m = n$, then the system is controllable.

It is clear that determining the controllability index is a delicate problem from the numerical view point because it is basically a rank-determination problem.

Fortunately, the block-Hessenberg pair $(H, \bar{B})$ of (A, B) not only determines if the pair (A, B) is controllable, but it also gives us the controllability index. In the block-Hessenberg pair $(H, \bar{B})$ in (6.7.2), k is the controllability index. Thus, for Example 6.7.2, the controllability index is 3.

Controllability Test in the Single-Input Case

In the single-input case, the controller-Hessenberg form of (A, b) becomes:

$$PAP^{\mathrm{T}} = H = \begin{pmatrix} h_{11} & h_{12} & \cdots & \cdots & h_{1n} \\ h_{21} & h_{22} & \cdots & \cdots & h_{2n} \\ 0 & h_{32} & \ddots & \cdots & h_{3n} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_{n,n-1} & h_{nn} \end{pmatrix}, \qquad Pb = \bar{b} = \begin{pmatrix} b_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{6.7.3}$$

Theorem 6.7.2. *(A, b) is controllable if the controller-Hessenberg pair $(H, \bar{b})$ is such that H is an unreduced upper Hessenberg, that is, $h_{i,i-1} \neq 0, i = 2, \ldots, n$, and $b_1 \neq 0$; otherwise, it is uncontrollable.*

We will give an independent proof of this test using the controllability criterion (ii) of Theorem 6.2.1.

Proof. Observe that $\mathrm{Rank}(b, Ab, \ldots, A^{n-1}b) = \mathrm{rank}(Pb, PAP^{\mathrm{T}}Pb, \ldots, PA^{n-1}P^{\mathrm{T}}Pb) = \mathrm{rank}(\bar{b}, H\bar{b}, \ldots, H^{n-1}\bar{b})$.

The last matrix is a lower triangular matrix with $b_1, h_{21}b_1, h_{21}h_{32}b_1, \ldots, h_{21}, \ldots, h_{n,n-1}b_1$ as the diagonal entries.

Since $h_{i,i-1} \neq 0$, $i = 2, \ldots, n$ and $b_1 \neq 0$, it follows that $\mathrm{rank}(b, Ab, \ldots, A^{n-1}b) = n$.

On the other hand, if any of $h_{i,i-1}$ or b_1 is zero, the matrix $(b, Ab, \ldots, A^{n-1}b)$ is rank deficient, and therefore, the system is uncontrollable. ■

Example 6.7.4 (*Example 6.6.2 Revisited*). Superiority of the algorithm over the other theoretical criteria.

To demonstrate the superiority of the test of controllability given by Theorem 6.7.2 over some of the theoretical criteria that we considered in the last section, Paige applied the controller-Hessenberg test to the same ill-conditioned problem as in Example 6.6.2. The computations gave $b_1 = 4.35887, h_{21} = 8.299699$,

$17 < \|h_{i,i-1}\| < 22, i = 3, 4, \ldots, 9$, and $h_{20,19} = 0.0000027$. Since $h_{20,19}$ is computationally zero in a single-precision computation, the system is uncontrollable, according to the test based on Theorem 6.7.2.

6.8 A NUMERICALLY EFFECTIVE TEST OF OBSERVABILITY

Analogous to the procedure of obtaining the form $(H, \bar{B})$ from (A, B), the pair (A, C) can be transformed to $(H, \bar{C})$, where

$$H = OAO^{\mathrm{T}} = \begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1k} \\ H_{21} & \ddots & & \vdots \\ & \ddots & \ddots & \vdots \\ 0 & & H_{k,k-1} & H_{kk} \end{pmatrix}, \tag{6.8.1}$$

$$\bar{C} = CO^{\mathrm{T}} = (0, C_1). \tag{6.8.2}$$

The pair (A, C) is observable if H is block unreduced (i.e., all the subdiagonal blocks have full rank) and the matrix C_1 has full rank.

The pair $(H, \bar{C})$ is said to be an **observer-Hessenberg pair**.

Flop-count: The construction of the observer-Hessenberg form this way requires roughly $6n^3 + 2n^2r$ flops.

Single-output case: In the single-output case, that is, when C is a row vector, the pair (A, C) is observable if H is an upper Hessenberg matrix and $\bar{C} = (0, \ldots, 0, c_1)$; $c_1 \neq 0$.

MATCONTROL note: MATCONTROL function **obserhs** can be used to obtain the reduction (6.8.1).

6.9 DISTANCE TO AN UNCONTROLLABLE SYSTEM

The concepts of controllability and observability are generic ones. Since determining if a system is controllable depends upon whether or not a certain matrix (or matrices) has full rank, it is immediately obvious from our discussion on numerical rank of a matrix in Chapter 4 that any uncontrollable system is arbitrary close to a controllable system. To illustrate this, let us consider the following well-known example (Eising 1984):

$$A = \begin{pmatrix} -1 & -1 & \cdot & \cdot & \cdot & \cdot & -1 & -1 \\ 1 & \cdot & & & & & & -1 \\ & 1 & \cdot & & & & & -1 \\ & & \cdot & \cdot & & & & \cdot \\ & & & \cdot & \cdot & & & \cdot \\ & & & & \cdot & \cdot & & \cdot \\ & 0 & & & & \cdot & \cdot & -1 \\ & & & & & & 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}. \tag{6.9.1}$$

The pair (A, b) is obviously controllable. However, it is easily verified that if we add $(-2^{1-n}, -2^{1-n}, \ldots, -2^{1-n})$ to the last row of (B, A), we obtain an uncontrollable system. Clearly, when n is large, the perturbation 2^{1-n} is small, implying that the original controllable system (A, B) is close to an uncontrollable system. **Thus, what is important in practice is knowledge of how close a controllable system is to an uncontrollable one rather than determining if a system is controllable or not.** To this end, we introduce, following Paige (1981), **a measure of the distance to uncontrollability,** denoted by $\mu(A, B)$:

$$\mu(A, B) \equiv \min\{\|\Delta A, \Delta B\|_2 \text{ such that the system defined by } (A + \Delta A, B + \Delta B) \text{ is uncontrollable}\}.$$

Here ΔA and ΔB are allowable perturbations over a field F. If the field F is $\mathbb{R}$, then we will use the symbol $\mu_{\mathrm{R}}(A, B)$ to distinguish it from $\mu(A, B)$.

The quantity $\mu(A, B)$ gives us a measure of the distance of a controllable pair (A, B) to the nearest uncontrollable pair. **If this distance is small, then the original controllable system is close to an uncontrollable system. If this distance is large, then the system is far from an uncontrollable system.**

Here is a well-known result on $\mu(A, B)$. See Miminis (1981), Eising (1984) and Kenney and Laub (1988). **Unless otherwise stated, the perturbations are assumed to be over the field of complex numbers, that is, $F = \mathbb{C}$.**

Theorem 6.9.1. *Singular Value Characterization to Distance to Uncontrollability.* $\mu(A, B) = \min \sigma_n(sI - A, B)$, *where* $\sigma_n(sI - A, B)$ *is the smallest singular value of* $(sI - A, B)$ *and* s *runs over all complex numbers.*

Proof. Suppose that $(A + \Delta A, B + \Delta B)$ is an uncontrollable pair. Then according to (v) of Theorem 6.2.1, we have

$$\operatorname{rank}(A + \Delta A - \lambda I, B + \Delta B) < n, \qquad \text{for some } \lambda \in \mathbb{C}.$$

Since the smallest perturbation that can make $\operatorname{rank}(A - \lambda I, B)$ less than n is $\sigma_n(A - \lambda I, B)$ (see Section 3.9.3 of Chapter 3), we have

$$\sigma_n(A - \lambda I, B) \leq \|\Delta A, \Delta B\|_2$$

and the equality holds if

$$(\Delta A, \Delta B) = -\sigma_n u_n v_n^*,$$

where σ_n is the smallest singular value of $(A - \lambda I, B)$, and u_n and v_n are the corresponding left and right singular vectors. Taking the minimum over all $\lambda \in \mathbb{C}$, and using criterion (v) of Theorem 6.2.1, we obtain the result. ∎

Algorithms for Computing $\mu(A, B)$

Based on Theorem 6.9.1, several algorithms (Miminis 1981; Eising 1984; Wicks and De Carlo 1991) have been developed in the last few years to compute $\mu(A, B)$ and $\mu_{\mathrm{R}}(A, B)$.

We will briefly describe here a Newton algorithm due to Elsner and He (1991), and an algorithm due to Wicks and DeCarlo (1991).

6.9.1 Newton's and the Bisection Methods for Computing the Distance to Uncontrollability

Let's denote $\sigma_n[sI - A, B]$ by $\sigma(s)$. The problem of finding $\mu(A, B)$ is then clearly the problem of minimizing $\sigma(s)$ over the complex plane.

To this end, define

$$f(s) = v_n^*(s)\begin{pmatrix} u_n(s) \\ 0 \end{pmatrix},$$

where $u_n(s)$and $v_n(s)$ are the normalized nth columns of U and V in the SVD of $(A - sI, B)$, that is, $(A - sI, B) = U\Sigma V^{\mathrm{T}}$. The function $f(s)$ plays an important role. The first and second derivatives of $\sigma(s) = \sigma(x + jy) = \sigma(x, y)$ can be calculated using this SVD. It can be shown that if $s = x + jy$, then

$$\frac{\partial \sigma}{\partial x} = \frac{\partial \sigma(x + jy)}{\partial x} = -\mathrm{Re} f(x+jy), \quad \text{and} \quad \frac{\partial \sigma}{\partial y} = \frac{\partial \sigma(x + jy)}{\partial y} = -\mathrm{Im}\, f(x+jy).$$

Knowing the first derivatives, the second derivatives can be easily calculated.

Hence the zeros of $f(s)$ are the critical points of the function $\sigma(s)$.

Thus, some well-established root-finding methods, such as **Newton's method**, or the **Bisection method** can be used to compute these critical points.

An interesting observation about the critical points is: **The critical points satisfy $s = u_n^*(s)Au_n(s)$, and hence they lie in the field of values of A.**

The result follows from the fact that $\sigma(s)f(s) = u_n^*(s)(A - sI)u_n(s)$, since $(A - sI, B)^*u_n(s) = \sigma(s)v_n(s)$. (For the definition of field of values, see Horn and Johnson (1985).)

To decide which critical points are local minima, one can use the following well-known criterion.

A critical point $s = x_c + jy_c$ of $\sigma(s)$ is a local minimum of $\sigma(x, y)$ if

$$\left(\frac{\partial^2 \sigma}{\partial x^2}\right)\left(\frac{\partial^2 \sigma}{\partial y^2}\right) - \left(\frac{\partial^2 \sigma}{\partial x \partial y}\right)^2 > 0 \quad \text{and} \quad \frac{\partial^2 \sigma}{\partial x^2} > 0.$$

Another sufficient condition is: If $\sigma_{n-1}(A - sI, B) > \sqrt{5}\sigma_n(A - sI, B)$, where $s = x_c + jy_c$ is a critical point, then (x_c, y_c) is a local minimum point of $\sigma(x, y)$.

Newton's method needs a starting approximation. The local minima of $\sigma(x, y)$, generally, are simple. Since all critical points s satisfy $u_n^*Au_n = s$, all minimum points $s = x + jy$ will lie in the field of values of A, and hence

$$\lambda_{\min}\left(\frac{A+A^{\mathrm{T}}}{2}\right) \le x \le \lambda_{\max}\left(\frac{A+A^{\mathrm{T}}}{2}\right) \quad \text{and} \quad \lambda_{\min}\left(\frac{A-A^{\mathrm{T}}}{2j}\right) \le y \le \lambda_{\max}\left(\frac{A-A^{\mathrm{T}}}{2j}\right),$$

where $\lambda_{\max}(C)$ and $\lambda_{\min}(C)$ denote the largest and smallest eigenvalues of the matrix C. Furthermore, since $\sigma_n(A - sI, B) = \sigma_n(A - \bar{s}I, B)$, the search for all local minimum points can be restricted to $0 \le y \le \lambda_{\max}((A - A^{\mathrm{T}})/2j)$.

Based on the above discussion, we now state Newton's algorithms for finding $\mu(A, B)$. Denote $x_k + jy_k$ by $\begin{pmatrix} x_k \\ y_k \end{pmatrix}$.

Algorithm 6.9.1. *Newton's Algorithm For Distance to Uncontrollability*
Inputs. *The matrices A and B.*
Output. *A local minimum of $\sigma(s)$.*
Step 0. *Choose $\binom{x_0}{y_0}$ using the above criterion.*
Step 1. *For $k = 0, 1, 2, \ldots$ do until convergence*

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \end{pmatrix} = \begin{pmatrix} x_k \\ y_k \end{pmatrix} - \theta_k \begin{pmatrix} p_{k_1} \\ p_{k_2} \end{pmatrix},$$

$$\textit{where } \begin{pmatrix} p_{k_1} \\ p_{k_2} \end{pmatrix} = \begin{pmatrix} \mathrm{Re}\,\dfrac{\partial f}{\partial x} & \mathrm{Re}\,\dfrac{\partial f}{\partial y} \\ \mathrm{Im}\,\dfrac{\partial f}{\partial x} & \mathrm{Im}\,\dfrac{\partial f}{\partial y} \end{pmatrix}^{-1} \begin{pmatrix} \mathrm{Re}\, f(x, y) \\ \mathrm{Im}\, f(x, y) \end{pmatrix},$$

choosing θ_k such that $\sigma(x_k - \theta_k p_{k1}, y_k - \theta_k p_{k2}) = \min\limits_{-1 \le \theta \le 1} \sigma(x_k - \theta p_{k1}, y_k - \theta p_{k2})$. (see Elsner and He (1991) for formulas for computing $\partial f/\partial x$ and $\partial f/\partial y$).

End.

Step 2. *If $s_C = \begin{pmatrix} x_f \\ y_f \end{pmatrix}$ is the final point upon conclusion of Step 2, then compute the smallest singular value σ_n of the matrix $(A - s_C I, B)$, and take σ_n as the local minimum of $\sigma(s)$.*

Choosing $\boldsymbol{\theta_k}$*: θ_k can be chosen using Newton's algorithm, again as follows: Define $g(\theta) = p_{k1}\mathrm{Re}\, f(\theta) + p_{k2}\mathrm{Im}\, f(\theta)$. Then $g'(\theta) = p_{k1}\mathrm{Re}\, f'(\theta) + p_{k2}\mathrm{Im}\, f'(\theta)$.*

Newton's Algorithm for Computing θ_k
Step 1. *Choose $\theta_0 = 1$.*

Step 2. *For $j = 1, 2, \ldots$ do until convergence*

$$\theta_{j+1} = \theta_j - \eta_j \frac{g(\theta_j)}{g'(\theta_j)},$$

where η_j is chosen such that

$$\sigma(x_k - \theta_{j+1} p_{k1},\ y_k - \theta_{j+1} p_{k2}) < \sigma(x_k - \theta_j p_{k1},\ y_k - \theta_j p_{k2})$$

End.

Remark

- Numerical experiments suggest that it is necessary to compute θ_k's only a few times to get a good initial point and then as soon as it becomes close 1, it can be set to 1. **Newton's method with $\theta_k = 1$ converges quadratically**.

Example 6.9.1 Elsner and He 1991. Let

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0.1 & 3 & 5 \\ 0 & -1 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 \\ 0.1 \\ 0 \end{pmatrix}.$$

$$\lambda_{\max}\left(\frac{A + A^{\mathrm{T}}}{2}\right) = 3.9925, \qquad \lambda_{\min}\left(\frac{A + A^{\mathrm{T}}}{2}\right) = -1.85133,$$

$$\lambda_{\max}\left(\frac{A - A^{\mathrm{T}}}{2j}\right) = 3.0745, \qquad \lambda_{\min}\left(\frac{A - A^{\mathrm{T}}}{2j}\right) = -3.0745.$$

Thus all zero points lie in the rectangular region given by $-1.8513 \leq x \leq 3.9925$, $-3.0745 \leq y \leq 3.0745$. Choose $x_0 = 1.5$ and $y_0 = 1$.

Then, $s_0 = 1.5 + j$.

$\theta_0 = 0.09935$, $\theta_1 = 0.5641$, $\theta_2 = 1.0012$. Starting from here, θ_k was set to 1.

The method converged in five steps. $s_{\mathrm{C}} = \binom{x_5}{y_5} = \binom{0.93708}{0.998571} = 0.93708 + 0.998571j$.

The minimum singular value of $(A - s_{\mathrm{C}} I, B) = 0.0392$.

Thus, $\mu(A, B) = 0.039238$.

MATLAB note: MATLAB codes for Algorithm 6.9.1 are available from the authors of the paper.

The Bisection Method (Real Case)

In the real case, the following bisection method can also be used to compute the zeros of $f(s)$.

Step 1. Find an interval $[a, b]$ such that $f(a)f(b) < 0$.

Step 2.

2.1. Compute $c = (a + b)/2$.

2.2. If $f(c)f(b) < 0$, then set $a = c$ and return to Step 2.1.
If $f(a)f(c) < 0$, then set $b = c$ and return to Step 2.1.
Step 3. Repeat Step 2 until c is an acceptable zero point of $f(s)$.
Note: For Example 6.9.1, $f(s)$ has only one real zero $s = 1.027337$, and $\mu_R(A, B) = 0.1725$.

6.9.2 The Wicks–DeCarlo Method for Distance to Uncontrollability

Newton's algorithm, described in Section 6.9.1 is based on minimization of $\sigma_n(sI - A, B)$ over all complex numbers s. It requires an SVD computation at each iteration.

In this section, we state an algorithm due to Wicks and DeCarlo (1991). The algorithm is also iterative in nature but "requires only two QR factorizations at each iteration without the need for searching or using a general minimization algorithm."

The algorithm is based on the following observation:

$$\mu(A, B) = \min_{u \in \mathbb{C}^n} ||(u^* A(I - uu^*)u^* B||, \tag{6.9.2}$$

subject to $u^* u = 1$. Based on this observation, they developed three algorithms for computing $\mu_R(A, B)$ and $\mu(A, B)$.

We state here why one of the algorithms (Algorithm II in Wicks and DeCarlo (1991)) is used for computing $\mu(A, B)$.

Definition 6.9.1. *Let the distance measure $d_1(A, B)$ be defined by*

$$\begin{aligned} [d_1(A, B)]^2 &= ||[e_n^*(A(I - e_n e_n^*)\ B)]||_2^2 \\ &= \sum_{j=1}^{n-1} |a_{nj}|^2 + \sum_{j=1}^{m} |b_{nj}|^2. \end{aligned}$$

Using the above notation, it has been shown in Wicks and DeCarlo (1991) that

$$\mu(A, B) = \min_{\substack{U \in C^{n \times n} \\ U^* U = I}} d_1(U^* A U, U^* B).$$

The algorithm proposed by them constructs a sequence of unitary matrices $U_1, U_2, \ldots$, such that

1. $A_{k+1} = U_k^* A_k U_k$, $B_{k+1} = U_k^* B$
2. $d_1(A_{k+1}, B_{k+1}) < d_1(A_k, B_k)$
3. $\lim_{k \to \infty} d_1(A_k, B_k)$ is a local minimum of (6.9.2).

Algorithm 6.9.2. *An Algorithm for Computing* $\mu(A, B)$
Inputs. *The matrices* A *and* B.
Output. $\mu(A, B)$.
Step 0. *Set* $A_1 \equiv A,\ B_1 \equiv B$.
Step 1. *For* $k = 1, 2, \ldots$ *until convergence.*
1.1. *Form* $M_k = (A_k - (a_{nn})_k I \quad B_k)$.
1.2. *Factor* $M_k = L_k V_k$, *where* L_k *is lower triangular and* V_k *is unitary.*
1.3. *Find the* QR *factorization of* $L_k = U_k^* R_k$.
1.4. *Set* $A_{k+1} = U_k^* A_k U_k, \quad B_{k+1} = U_k^* B_k$.
1.5. *If* $d_1(A_{k+1}, B_{k+1}) = d_1(A_k, B_k)$, *stop.*
End.

Proof. The proof amounts to showing that $d_1(A_k, B_k) \geq |r_{nn}|$, where

$$R_k = \begin{pmatrix} r_{11} & \cdots & \cdots & r_{1n} & \\ 0 & r_{22} & & r_{2n} & 0 \\ \vdots & & \ddots & \vdots & \\ 0 & & & r_{nn} & \end{pmatrix}$$

and as such $|r_{nn}| = d_1(U_k^* A_k U_k, U_k^* B_k)$. ■

For details, the readers are referred to Wicks and DeCarlo (1991).

Example 6.9.2.

$$A = \begin{pmatrix} 0.950 & 0.891 & 0.821 & 0.922 \\ 0.231 & 0.762 & 0.445 & 0.738 \\ 0.607 & 0.456 & 0.615 & 0.176 \\ 0.486 & 0.019 & 0.792 & 0.406 \end{pmatrix},$$

$$B = \begin{pmatrix} 0.9350 & 0.0580 & 0.1390 \\ 0.9170 & 0.3530 & 0.2030 \\ 0.4100 & 0.8130 & 0.1990 \\ 0.8940 & 0.0100 & 0.6040 \end{pmatrix}.$$

Let the tolerance for stopping the iteration be: Tol $= 0.00001$.
Define $\mu_k = d_1(A_k, B_k)$.

The algorithm produces the following converging sequence of μ_k:

k	μ_k	k	μ_k
0	1.42406916966838	10	0.41450782001833
1	0.80536738314449	11	0.41450781529413
2	0.74734006994998	12	0.41450781480559
3	0.52693889988172	13	0.41450781475487
4	0.42241562062172	14	0.41450781474959
5	0.41511102322896	15	0.41450781474904
6	0.41456112538077	16	0.41450781474899
7	0.41451290008455	17	0.41450781474898
8	0.41450831981698	18	0.41450781474898
9	0.41450786602577	19	0.41450781474898

After 19 iterations the algorithm returns $\mu = 0.41450781474898$.

MATCONTROL note: Algorithm 6.9.2 has been implemented in MATCONTROL function **discntrl**.

6.9.3 A Global Minimum Search Algorithm

The algorithms by Elsner and He (1991) and Wicks and DeCarlo (1991) are guaranteed only to converge to a local minimum rather than a global minimum. A global minimum search algorithm was given by Gao and Neumann (1993). Their algorithm is based on the observation that if $\text{rank}(B) < n$, then the minimization problem can be transformed to a minimization problem in the bounded region $\{(x, z) | x \leq \|A\|_2, |z| \leq \|A\|_2\}$ in the two-dimensional real plane.

The algorithm then progressively partitions this region into simplexes and finds lower and upper bounds for $\mu(A, B)$ by determining if the vertices (x_k, z_k) satisfy

$$z_k > \min_{y \in \mathbb{R}} \sigma_{\min}(A - (x_k + jy)I, B).$$

These bounds are close to each other if $\mu(A, B)$ is small. "If $\mu(A, B)$ is not small, then the algorithm produces a lower bound which is not small, thus leading us to a safe conclusion that (A, B) is not controllable."

For details of the algorithm, see Gao and Neumann (1993). See also **Exercise 6.26**.

6.10 DISTANCE TO UNCONTROLLABILITY AND THE SINGULAR VALUES OF THE CONTROLLABILITY MATRIX

Since the rank of the controllability matrix C_M determines whether a system is controllable or not, and the most numerically effective way to determine the rank of a matrix is via the singular values of the matrix, it is natural to wonder, what roles do the singular values of the controllability matrix play in deciding if a given controllable system is near an uncontrollable system. (Note that Theorem 6.9.1 and the associated algorithm for computing $\mu(A, B)$ use the singular values of $(A - sI, B)$).

The following result due to Boley and Lu (1986) sheds some light in that direction. We state the result without proof. Proof can be found in Boley and Lu (1986).

Theorem 6.10.1. *Let (A, B) be a controllable pair. Then,*

$$\mu(A, B) \leq \mu_R(A, B) \leq \left(1 + \frac{\|C_p\|}{\sigma_{n-1}}\right)\sigma_n,$$

where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{n-1} \geq \sigma_n$ are the singular values of the controllability matrix $C_M = (B, AB, \ldots, A^{n-1}B)$ and C_p is a companion matrix for A.

Example 6.10.1. We consider Example 6.9.1 again.

The singular values of the controllability matrix are 2.2221, 0.3971, 0.0227.

The companion matrix C_p is calculated as follows:

$$x_1 = (1, 0, 0)^T, \qquad x_2 = Ax_1 = (1, 0.1, 0)^T, \qquad x_3 = Ax_2 = (1.1, 0.4, -0.1)^T.$$

Then the matrix $X = (x_1, x_2, x_3)$ is such that

$$X^{-1}AX = C_p = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & -3.9 \\ 0 & 1 & 3 \end{pmatrix}.$$

According to Theorem 6.10.1, we then have

$$\mu(A, B) \leq \mu_R(A, B) \leq \left(1 + \frac{5.3919}{0.3971}\right) \times 0.02227 = 0.3309.$$

Remark

- The above theorem can be used to predict the order of perturbations needed to transform a controllable system to an uncontrollable system. *It is the largest gap between the consecutive singular values.* (**However, note that, in general, the singular values of the controllability matrix cannot be used directly to make a prediction of how close the system is to an uncontrollable system.**)

In other words, it is **not** true that one can obtain a nearly uncontrollable system by applying perturbations ΔA, ΔB, with norm bounded by the smallest nonzero singular value of the controllability matrix.

6.11 SOME SELECTED SOFTWARE

6.11.1 MATLAB Control System Toolbox

canon—State-space canonical forms.
ctrb, obsv—Controllability and observability matrices.
gram—Controllability and observability gramians.
ctrbf—Controllability staircase form.
obsvf—observability staircase form.

6.11.2 MATCONTROL

CNTRLHS—Finding the controller-Hessenberg form.
CNTRLHST—Finding the Controller-Hessenberg form with triangular subdiagonal blocks.
OBSERHS—Finding the observer-Hessenberg form.
CNTRLC—Find the controller-canonical form (Lower Companion).
DISCNTRL—Distance to controllability using the Wicks–DeCarlo algorithm.

6.11.3 CSP-ANM

- **Reduction to controller-Hessenberg and observer-Hessenberg forms**
  - Block controller-Hessenberg forms are computed by `controller-HessenbergForm` [*system*] and `LowercontrollerHessen-bergForm` [*system*].
  - Block observer-Hessenberg forms are computed by `Observer-HessenbergForm` [*system*] and `UpperObserverHessen-bergForm` [*system*].
- **Controllability and observability tests**
  - Tests of controllability and observability using block controller-Hessenberg and block observer-Hessenberg forms are performed via `Controllable` [*system*, `ControllabilityTest` → `Full-RankcontrollerHessenbergBlocks`] and `Observable` [*system*, `ObservabilityTest` → `FullRankObserver-HessenbergBlocks`].
  - Tests of controllability and observability of a stable system via positive definiteness of Gramians are performed via `Controllable` [*system*, `ControllabilityTest` → `PositiveDiagonal-CholeskyFactorControllabilityGramian`] and `Observable` [*system*, `ObservabilityTest` → `Positive-DiagonalCholeskyFactorObservabilityGramian`].

6.11.4 SLICOT

Canonical and quasi canonical forms:

AB01MD—Orthogonal controllability form for single-input system
AB01ND—Orthogonal controllability staircase form for multi-input system
AB01OD—Staircase form for multi-input system using orthogonal transformations
TB01MD—Upper/lower controller-Hessenberg form
TB01ND—Upper/lower observer-Hessenberg form
TB01PD—Minimal, controllable or observable block Hessenberg realization
TB01UD—Controllable block Hessenberg realization for a state-space representation
TB01ZD—Controllable realization for single-input systems

6.11.5 MATRIX$_X$

Purpose: Obtain controllable part of a dynamic system.
Syntax: [SC, NSC, T]= CNTRLABLE (S, NS, TOL)

Purpose: Compute observable part of a system.
Syntax: [SOBS, NSOBS, T]= OBSERVABLE (S, NS, TOL)

Purpose: Staircase form of a system matrix.
Syntax: [SST, T, NCO]= STAIR (S, NS, TOL)

6.12 SUMMARY AND REVIEW

Algebraic Criteria of Controllability and Observability

Controllability and observability are two most fundamental concepts in control theory. The algebraic criteria of controllability and observability are summarized in **Theorems 6.2.1** and **6.3.1**, respectively.

Unfortunately, these algebraic criteria very often do not yield numerically viable tests for controllability and observability. The numerical difficulties with these criteria as practical tests of controllability are discussed and illustrated in Section 6.6. The pair (A, B) in Example 6.6.1 is a controllable pair; however, it is shown that the Example 6.6.1 is a controllable pair; however, it is shown that criterion (ii) of Theorem 6.2.1 leads to an erroneous conclusion due to a computationally small singular value of the controllability matrix. Similarly, in Example 6.6.2, it is shown how an obviously uncontrollable pair can be taken as a controllable pair by using the eigenvalue criterion of controllability (Criterion (v) of Theorem 6.2.1) as a numerical test, due to the ill-conditioning of the eigenvalues of the matrix A.

Numerically Effective Tests of Controllability and Observability

Computationally viable tests of controllability and observability are given in **Sections 6.7** and **6.8**. These tests are based on the reductions of the pairs (A, B) and (A, C), respectively, to **controller-Hessenberg** and **observer-Hessenberg** forms. These forms can be obtained by using orthogonal transformations and the tests can be shown to be numerically stable.

Indeed, when controller-Hessenberg test is applied to Example 6.6.2, it was concluded correctly that, in spite of the ill-conditioning of the eigenvalues of A, the pair (A, B) is uncontrollable.

Distance to Uncontrollability

Since determining the rank of a matrix is numerically a delicate problem and the problem is sensitive to small perturbations, **in practice, it is more important to find when a controllable system is close to an uncontrollable system.** To this end, a practical measure of the distance to uncontrollability, denoted by $\mu(A, B)$, is introduced in Section 6.9: $\mu(A, B) = \min\{\|\Delta A, \Delta B\|_2$ such that the pair $(A + \Delta A, B + \Delta B)$ is controllable.

A well-known characterization of $\mu(A, B)$ is given in **Theorem 6.9.1.** This theorem states: $\mu(A, B) = \min \sigma_n(sI - A, B)$, where $\sigma_n(sI - A, B)$ is the smallest singular value of the matrix $(sI - A, B)$ and s runs over all complex numbers.

Two algorithms (**Algorithms 6.9.1** and **6.9.2**), have been described to compute $\mu(A, B)$.

6.13 CHAPTER NOTES AND FURTHER READING

Controllability and observability are two most basic concepts in control theory. The results related to controllability and observability can be found in any standard book on linear systems (e.g., Kalman *et al.* 1969; Brockett 1970; Rosenbrock 1970; Luenberger 1979; Kailath 1980; Chen 1984; DeCarlo 1989; Brogan 1991; etc.).

For details on the staircase algorithms for finding the controller-Hessenberg and observer-Hessenberg forms, see Boley (1981), Paige (1981), Van Dooren and Verhaegen (1985), etc. For computation of the Kalman decomposition, see Boley (1980, 1991), etc. For more on the concept of the distance to uncontrollability and related algorithms, see Boley (1987), Boley and Lu (1986), Eising (1984), Wicks and DeCarlo (1991), Elsner and He (1991), Paige (1981), Miminis (1981), Kenney and Laub (1988), and Gao and Neumann (1993). For a test of controllability via real Schur form, see Varga (1979).

Exercises

6.1 Prove that (A, B) is controllable if and only if for a constant matrix F, the matrix $(A + BF, B)$ is controllable, that is, the controllability of a system does not change under state feedback. **(The concept of state feedback is defined in Chapter 10.)**

6.2 Construct an example to show that the observability of a system may change under state feedback.

6.3 A matrix A is called a **cyclic matrix** if in the JCF of A, there is only one Jordan box associated with each distinct eigenvalue.
Let A be a cyclic matrix and let the pair (A, B) be controllable. Then prove that for almost all vectors v, the pair (A, Bv) is controllable.

6.4 Give a 2×2 example to show that the cyclicity assumption is essential for the result of Problem 6.3 to hold.

6.5 Show that (A, c) is observable if and only if there exists a vector k such that

$$\left(\begin{pmatrix} c \\ k \end{pmatrix}, \ A - bk\right)$$

is observable.

6.6 Prove the parts (i), (ii), (iv)–(vi) of Theorem 6.3.1.

6.7 Prove Theorem 6.4.2.

6.8 Using Theorems 6.4.1 and 6.4.2, give a proof of Theorem 6.4.3 (**The Kalman Canonical Decomposition Theorem**).

6.9 Prove that the change of variable $\bar{x} = Tx$, where T is nonsingular, preserves the controllability and observability of the system (A, B, C).

6.10 Work out an algorithm to compute the nonsingular transforming matrix that transforms the pair (A, b) to the upper companion form. When can the matrix transforming T be highly ill-conditioned? Construct a numerical example to support your statement.

6.11 Apply the test based on the eigenvalue criterion of controllability to Example 6.6.2 and show that this test will do better than the one based on the criterion (ii) of Theorem 6.2.1.

6.12 Applying the staircase algorithm to the pair (A, b) in Example 6.6.2, show that the pair (A, b) is uncontrollable.

6.13 If the controller-Hessenberg pair $(H, \bar{B})$ of the controllable system (A, B) is such that the subdiagonal blocks of H are nearly rank-deficient, then the system may be very near to an uncontrollable system.

(a) Construct examples both in the single- and multi-input cases in support of the above statement.

(b) Construct another example to show that the converse is not necessarily true, that is, even if the subdiagonal blocks of H have robust ranks, the system may be close to an uncontrollable system.

6.14 Show that to check the controllability for the pair (A, B), where

$$A = \mathrm{diag}(1, 2^{-1}, \ldots, 2^{1-n}) \quad \text{and} \quad B = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix},$$

the eigenvalue-criterion for controllability (the PBH criterion) will do better than the criterion (ii) of Theorem 6.2.1.

6.15 Let $(H = (H_{ij}), \bar{B})$ be the controller-Hessenberg form of the pair (A, B). Then prove the following:

(a) If $H_{i,i-1} = 0$ for any i, then (A, B) is not controllable.

(b) $\mu(A, B) \leq \min\limits_{1 \leq i \leq k} \|H_{i,i-1}\|_2$.

6.16 Prove that $\mu(A, B)$ remains invariant under an orthogonal transformation.

6.17 Let $(\tilde{A}, \tilde{b})$ be as in (6.5.2). Prove that $\mu(\tilde{A}, \tilde{b}) \leq \sin(\pi/n)$ (Kenney and Laub 1988).

6.18 Develop an algorithm for the reduction of the pair (A, C) to the observer-Hessenberg form (6.8.1), without invoking the algorithm for the controller-Hessenberg reduction to the pair $(A^{\mathrm{T}}, C^{\mathrm{T}})$. How can one obtain the observability indices of the pair (A, C) from this form?

6.19 Construct a simple example to show that the minimum which yields $\mu(A, B)$ is not achieved when s is an eigenvalue of A.

6.20 Rework Example 6.9.1 with the initial point as one of the eigenvalues of $F = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$, where (C, D) is a random matrix such that F is square.

6.21 Apply the bisection method to Example 6.9.1 to find an estimate of $\mu_R(A, B)$.

6.22 Find $\mu(A, B)$ and $\mu_R(A, B)$, where A and B are given by:

$$A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & \ddots & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}_{10\times 10} \quad \text{and} \quad B = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}_{10\times 1}.$$

6.23 Derive Newton's algorithm for computing $\mu_R(A, B)$.

6.24 (Laub and Linnemann 1986). Consider the controllable pair

$$(H, \bar{b}) = \left(\begin{pmatrix} -4 & 0 & 0 & 0 \\ \alpha & -3 & 0 & 0 \\ 0 & \alpha & -2 & 0 \\ 0 & 0 & \alpha & -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right),$$

with $0 < \alpha \leq 1$. Show (experimentally or mathematically) that the pair $(H, \bar{b})$ is close to an uncontrollable pair.

6.25 Let (A, B) be controllable. Let $B = (B_1, B_2)$, with B_1 consisting of minimum number of inputs such that (A, B) is controllable. Then prove that (A, B_1) is closer to an uncontrollable system than (A, B) is; that is, prove that

$$\min \sigma_n(B_1, A - sI) \leq \min \sigma_n(B, A - sI), \quad s \in \mathbb{C}$$

6.26 (Gao and Neumann 1993). Let $\lambda_0 \in \mathbb{C}$ and let $p \in \mathbb{C}$ be on the unit circle. Consider the straight line $\lambda = \lambda_0 + tp$, $t \in \mathbb{R}$. Then prove that

$$\min_{t \in \mathbb{R}} \sigma_{\min}(A - (\lambda_0 + tp)I, B) \leq \alpha$$

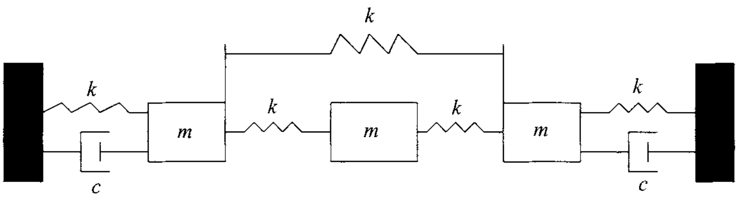
if and only if the matrix

$$G(\alpha) = \begin{pmatrix} \bar{p}(A - \lambda_0 I) & \bar{p}(BB^* - \alpha^2 I) \\ -pI & p(A^* - \bar{\lambda}_0 I) \end{pmatrix}$$

has a real eigenvalue.

Based on the above result, derive an algorithm for computing $\mu_{\mathrm{R}}(A, B)$.
(Hint: take $\lambda_0 = 0,\ p = 1$.)
Test your algorithm with Example 6.9.1.

6.27 (a) Construct a state-space representation of the following second-order model.



$$\begin{bmatrix} m & & \\ & m & \\ & & m \end{bmatrix}\begin{pmatrix} \ddot{u}_1 \\ \ddot{u}_2 \\ \ddot{u}_3 \end{pmatrix} + \begin{bmatrix} c & & \\ & 0 & \\ & & c \end{bmatrix}\begin{pmatrix} \dot{u}_1 \\ \dot{u}_2 \\ \dot{u}_3 \end{pmatrix}$$
$$+ \begin{bmatrix} 3k & -k & -k \\ -k & 2k & -k \\ -k & -k & 3k \end{bmatrix}\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

(b) Show that the system is not controllable for

$$b = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad b = \begin{pmatrix} \alpha \\ \beta \\ \alpha \end{pmatrix}, \quad \text{or} \quad b = \begin{pmatrix} \beta \\ 0 \\ -\beta \end{pmatrix}.$$

6.28 Consider Example 5.2.6 on the motion of an orbiting satellite with $d_0 = 1$. Let $x(t) = (x_1(t), x_2(t), x_3(t), x_4(t))^{\mathrm{T}}$, $u(t) = (u_1(t), u_2(t))^{\mathrm{T}}$, and $y(t) = (y_1(t), y_2(t))^{\mathrm{T}}$.
(a) Show that one of the states cannot be controlled by the radial force $u_1(t)$ alone, but all the states can be controlled using the tangential force $u_2(t)$.
(b) Show that all the states are observable using both the outputs; however, one of the states cannot be observed by $y_1(t)$ alone.

6.29 (Boley 1985). Let $(H, \bar{b})$ be the controller-Hessenberg pair of the controllable pair (A, b) such that $\|H\|_2 + \|\bar{b}\|_2 \leq \frac{1}{4}$. Then prove that the quantity $|\bar{b}_1 h_{21} h_{32} \cdots h_{n,n-1}|$ gives a lower bound on the perturbations needed to obtain an uncontrollable pair. Construct an example to support this.

6.30 Does the result of the preceding exercise hold in the multi-input case? Prove or disprove.

6.31 Consider the example of balancing a stick on your hand (Example 5.2.4). We know from our experience that a stick can be balanced. Verify this using a criterion of controllability. Take $L = 1$.

$$A = \begin{pmatrix} 0 & 1 \\ \frac{g}{L} & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ -\frac{g}{L} \end{pmatrix}.$$

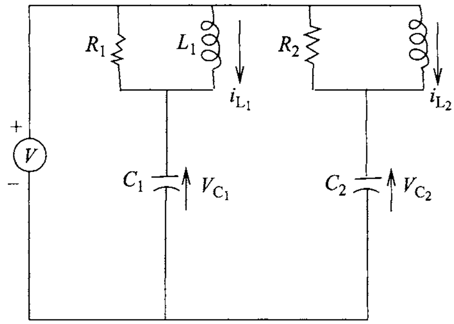


FIGURE 6.1: Uncontrollability of an electrical network.

6.32 **(An Uncontrollable System)** (Szidarovszky and Bahill (1991, pp. 223–224)). Consider the electric network in Figure 6.1 with two identical circuits in parallel. Intuitively, it is clear that there cannot exist a single input that will bring one circuit to one state and the other to a different state. Verify this using a criterion of controllability. Take $L_1 = L_2 = 1$, $C_1 = C_2 = 1$, and $R_1 = R_2 = 1$.

References

Boley D. "On Kalman's procedure for the computation of the controllable/observable canonical form," *SIAM J. Control Optimiz.*, Vol. 18, pp. 624–626, 1980.

Boley D. *Computing the Controllability/Observability of a Linear Time-Invariant Dynamic System: A Numerical Approach*, Ph.D. thesis, Computer Science Department, Stanford University, Stanford, CA, June 1981.

Boley D. and Lu W.-S. "Measuring how far a controllable system is from an uncontrollable system," *IEEE Trans. Autom. Control*, Vol. AC-31, pp. 249–251, 1986.

Boley D. A. "Perturbation result for linear control problems," *SIAM J. Alg. Discr. Methods*, Vol. 6, pp. 66–72, 1985.

Boley D. "Computing the Kalman decomposition, an optimal method," *IEEE Trans. Autom. Control*, Vol. AC-29, pp. 51–53, 1984; correction in Vol. 36, p. 1341, 1991.

Brockett R. *Finite Dimensional Linear Systems*, Wiley, New York, 1970.

Brogan W.L. *Modern Control Theory*, 3rd edn, Prentice Hall, Englewood Cliffs, NJ, 1991.

Chen C.-T. *Linear Systems Theory and Design*, CBS College Publishing, New York, 1984.

Datta B.N. *Numerical Linear Algebra and Applications*, Brooks/Cole Publishing Company, Pacific Grove, CA, 1995.

DeCarlo R.A. *Linear Systems: A State Variable Approach with Numerical Implementation*, Prentice Hall, Englewood Cliffs, NJ, 1989.

Eising R. "Between controllable and uncontrollable," *Syst. Control Lett.*, Vol. 4, pp. 263–264, 1984.

Elsner L. and He C. "An algorithm for computing the distance to uncontrollability," *Syst. Control Lett.*, Vol. 17, pp. 453–464, 1991.

Gao M. and Neumann M. "A global minimum search algorithm for estimating the distance to uncontrollability," *Lin. Alg. Appl.*, Vol. 188/189, pp. 305–350, 1993.

Golub G.H. and Van Loan C.F. *Matrix Computations*, 3rd edn, Johns Hopkins University, Baltimore, MD, 1996.

Hautus M.L.J. "Controllability and observability conditions of linear autonomous systems," *Proc. Kon. Ned. Akad. Wetensh*, Ser. A. Vol. 72, pp. 443–448, 1969.

Horn R.A. and Johnson C. *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.

Kailath T. *Linear Systems*, Prentice Hall, Englewood Cliffs, NJ, 1980.

Kalman R.E., Flab P.L., and Arbib M.A. *Topics in Mathematical Systems Theory*, McGraw Hill, New York, 1969.

Kenney C.S. and Laub A.J. "Controllability and stability radii for companion form systems," *Math. Control, Signals Syst.*, Vol. 1, pp. 239–256, 1988.

Laub A.J. and Linnemann A. "Hessenberg and Hessenberg/triangular forms in linear systems theory," *Int. J. Control*, Vol. 44, pp. 1523–1547, 1986b.

Luenberger D.G. *Introduction to Dynamic Systems: Theory, Models & Applications*, John Wiley & Sons, New York, 1979.

Miminis G.S. *Numerical Algorithms for Controllability and Eigenvalue Allocations*, Master Thesis, School of Computer Science, McGill University, Montreal, Canada, May, 1981.

Paige C.C. "Properties of numerical algorithms related to computing controllability," *IEEE Trans. Autom. Control*, Vol. AC-26, pp. 130–138, 1981.

Patel R. "Computation of Minimal order state-space realization and observability indices using orthogonal transformation," *Int. J. Control.*, Vol. 33, pp. 227–246, 1981.

Rosenbrock M.M. *State-Space and Multivariable Theory*, John Wiley, New York, 1970.

Szidarovszky F. and Terry Bahill A. *Linear Systems Theory*, CRC Press, Boca Raton, 1991.

Van Dooren P.M. and Verhaegen M. "On the use of unitary state-space transformations, Contemporary Mathematics," in *Contemporary Mathematics Series* (Brualdi R. *et al.*, eds.), Vol. 47, pp. 447–463, Amer. Math. Soc., Providence, RI, 1985.

Varga A. "Numerically reliable algorithm to test controllability," *Electronic Lett.*, Vol. 15, pp. 452–453, 1979.

Wicks M. and DeCarlo A. "Computing the distance to an uncontrollable system," *IEEE Trans. Autom. Control*, Vol. 36, pp. 39–49, 1991.