Chapter 3

Controllability

All the material in this chapter pertains to the differential equation

$$\dot{x} = Ax + Bu, , \qquad x(0) = 0 , \tag{3.1}$$

where $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$, so $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$. We study the concept of controllability and two important properties of controllability: invariance under change of basis and invariance under state feedback. We also discuss controllable canonical form for single input systems, which is very useful for the pole assignment problem discussed in the next chapter.

3.1 Reachable States

Consider the following problem. For fixed t_1 and a given vector v in $\mathbb{R}^n$, does there exist a control input u, defined on $[0, t_1]$, such that the solution of (3.1) satisfies $x(t_1) = v$? We shall refer to this problem as the *reachability problem.*

Let $\mathcal{U}$ be the space of piecewise continuous functions with finite energy in every finite time interval. Define the linear operator $L_c : \mathcal{U} \to \mathbb{R}^n$ by

$$L_c u = \int_0^{t_1} e^{A(t_1-s)} B u(s) ds .$$

The reachability problem is equivalent to the solvability of the linear equation

$$L_c u = v .$$

To bring out the analogy a bit further, recall that the linear equation

$$A\xi = b$$

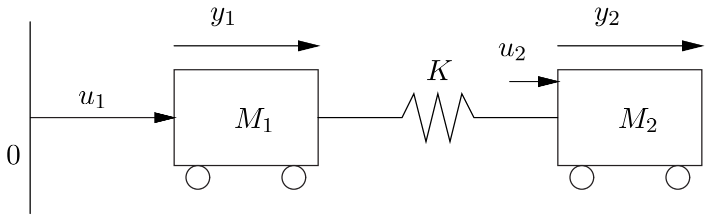
has a solution if and only if $b \in \mathcal{R}(A)$ where $\mathcal{R}(A)$ is the range of the matrix A. Similarly, the reachability problem is solvable if and only if $v \in \mathcal{R}(L_c)$. Then we say that the state v is reachable at time t_1. The set of reachable states is given by $\mathcal{R}(L_c)$. Every state is reachable iff L_c is onto. If every state is reachable then we say that the system (A, B) is *controllable.*

Define the *controllability matrix* Q_c, which has dimension $n \times nm$, as

$$Q_c = [\ B \quad AB \quad A^2B \quad \ldots \quad A^{n-1}B\ ] .$$

We will see shortly that the rank of this matrix determines whether a system is controllable.

Example 3.1.1.



The system equations are

$$\begin{aligned} M_1\ddot{y}_1 &= -K(y_1 - y_2) + u_1 \\ M_2\ddot{y}_2 &= -K(y_2 - y_1) + u_2\,. \end{aligned}$$

We define the state vector

$$x = \begin{bmatrix} y_1 \\ \dot{y}_1 \\ y_2 \\ \dot{y}_2 \end{bmatrix}.$$

Then the state equations are

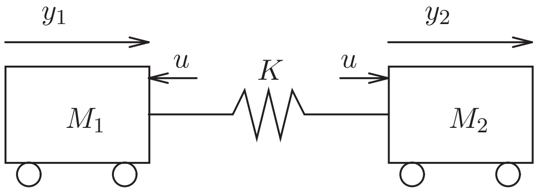
$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{-K}{M_1} & 0 & \frac{K}{M_1} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{K}{M_2} & 0 & \frac{-K}{M_2} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \frac{1}{M_1} & 0 \\ 0 & 0 \\ 0 & \frac{1}{M_2} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

The controllability matrix is

$$Q_c \;=\; \begin{bmatrix} 0 & 0 & \frac{1}{M_1} & 0 & 0 & 0 & \frac{-K}{M_1^2} & \frac{K}{M_1 M_2} \\ \frac{1}{M_1} & 0 & 0 & 0 & \frac{-K}{M_1^2} & \frac{K}{M_1 M_2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{M_2} & 0 & 0 & \frac{K}{M_1 M_2} & \frac{-K}{M_2^2} \\ 0 & \frac{1}{M_2} & 0 & 0 & \frac{K}{M_1 M_2} & \frac{-K}{M_1^2} & 0 & 0 \end{bmatrix}$$

which has 4 linearly independent columns so that it is full rank.

Example 3.1.2.



The system equations are

$$\begin{aligned} M_1\ddot{y}_1 &= -u - K(y_1 - y_2) \\ M_2\ddot{y}_2 &= u - K(y_2 - y_1) \end{aligned}$$

and defining the state vector

$$x = \begin{bmatrix} y_1 \\ \dot{y}_1 \\ y_2 \\ \dot{y}_2 \end{bmatrix},$$

we obtain the state equations

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \frac{-K}{M_1} & 0 & \frac{K}{M_1} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{K}{M_2} & 0 & \frac{-K}{M_2} & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ -\frac{1}{M_1} \\ 0 \\ \frac{1}{M_2} \end{bmatrix} u \, .$$

The controllability matrix is

$$Q_c = \begin{bmatrix} 0 & -\frac{1}{M_1} & 0 & \frac{K}{M_1^2} + \frac{K}{M_1 M_2} \\ -\frac{1}{M_1} & 0 & \frac{K}{M_1^2} + \frac{K}{M_1 M_2} & 0 \\ 0 & \frac{1}{M_2} & 0 & -\frac{K}{M_1 M_2} - \frac{K}{M_2^2} \\ \frac{1}{M_2} & 0 & -\frac{K}{M_1 M_2} - \frac{K}{M_2^2} & 0 \end{bmatrix} .$$

Note that for this Q_c

$$\begin{aligned} \text{the 3rd column} \quad &= \quad \begin{bmatrix} 0 \\ \frac{1}{M_1}\left(\frac{K}{M_1} + \frac{K}{M_2}\right) \\ 0 \\ \frac{1}{M_2} \end{bmatrix} \\ &= \quad -\left(\frac{K}{M_1} + \frac{K}{M_2}\right) \begin{bmatrix} \text{1st} \\ \text{column} \end{bmatrix} \end{aligned}$$

so that only two columns in Q_c are linearly independent.

Associated with Q_c, we define the *controllable subspace* which is $\mathcal{R}(Q_c)$. Note that for the two cart-two force system $\mathcal{R}(Q_c) = \mathbb{R}^4$ while for the two cart-one force system $\mathcal{R}(Q_c) = \mathbb{R}^2$.

The infinite dimensional vector space $\mathcal{U}$ has an inner product

$$\langle u, w \rangle_{\mathcal{U}} := \int_0^{t_1} u(\tau)^T w(\tau) d\tau \, .$$

Referring to the Appendix, the adjoint operator $L_c^* : \mathbb{R}^n \rightarrow \mathcal{U}$ of L_c is

$$(L_c^* v)(\tau) = B^T e^{(t_1 - \tau)A^T} v \, .$$

We define, for each $t > 0$, the *controllability gramian*

$$\begin{aligned} W_c(t) := (L_c L_c^*)(t) \quad &= \quad \int_0^t e^{A(t-\tau)} B B^T e^{A^T(t-\tau)} d\tau \\ &= \quad \int_0^t e^{A\tau} B B^T e^{A^T \tau} d\tau \, . \end{aligned}$$

We will shortly make use of the fact that $\mathcal{R}(L_c) = \mathcal{R}(L_c L_c^*)$; the proof is in the Appendix. This result is useful because $L_c L_c^* : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a matrix.

Example 3.1.3.

Suppose we have

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

$$\begin{aligned}
(sI-A)^{-1} &= \begin{bmatrix} s+1 & -1 \\ 0 & s+2 \end{bmatrix}^{-1} = \frac{\begin{bmatrix} s+2 & 1 \\ 0 & s+1 \end{bmatrix}}{(s+1)(s+2)} \\
&= \begin{bmatrix} \frac{1}{s+1} & \frac{1}{(s+1)(s+2)} \\ 0 & \frac{1}{s+2} \end{bmatrix} = \begin{bmatrix} \frac{1}{s+1} & \frac{1}{s+1} - \frac{1}{s+2} \\ 0 & \frac{1}{s+2} \end{bmatrix} \\
e^{At} &= \mathcal{L}^{-1}(sI-A)^{-1} = \begin{bmatrix} e^{-t} & e^{-t} - e^{-2t} \\ 0 & e^{-2t} \end{bmatrix} \\
W_c(t) &= \textstyle\int_0^t \begin{bmatrix} e^{-\tau} & e^{-\tau} - e^{-2\tau} \\ 0 & e^{-2\tau} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} [0 \;\; 1] \begin{bmatrix} e^{-\tau} & 0 \\ e^{-\tau} - e^{-2\tau} & e^{-2\tau} \end{bmatrix} d\tau \\
&= \textstyle\int_0^t \begin{bmatrix} e^{-\tau} & e^{-\tau} - e^{-2\tau} \\ 0 & e^{-2\tau} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ e^{-\tau} - e^{-2\tau} & e^{-2\tau} \end{bmatrix} d\tau \\
&= \textstyle\int_0^t \begin{bmatrix} e^{-2\tau} - 2e^{-3\tau} + e^{-4\tau} & e^{-3\tau} - e^{-4\tau} \\ e^{-3\tau} - e^{-4\tau} & e^{-4\tau} \end{bmatrix} d\tau \\
&= \begin{bmatrix} \frac{1}{2}(1-e^{-2t}) - \frac{2}{3}(1-e^{-3t}) + \frac{1}{4}(1-e^{-4t}) & \frac{1}{3}(1-e^{-3t}) - \frac{1}{4}(1-e^{-4t}) \\ \frac{1}{3}(1-e^{-3t}) - \frac{1}{4}(1-e^{-4t}) & \frac{1}{4}(1-e^{-4t}) \end{bmatrix}.
\end{aligned}$$

Note that W_c is a symmetric $n \times n$ matrix.

We want to find a characterization of $\mathcal{R}(L_c)$ that allows us to determine in a computationally direct way whether a linear system is controllable. The main result of this section is the following.

Theorem 3.1.1. *$\mathcal{R}(L_c) = \mathcal{R}(Q_c)$.*

Proof. First we will show that $\mathcal{R}(L_c) \subset \mathcal{R}(Q_c)$. The Cayley-Hamilton theorem (refer to the Appendix for a proof) says that a matrix satisfies its own characteristic polynomial; that is, if $p(s) = s^n + a_{n-1}s^{n-1} + \cdots + a_0$ is the characteristic polynomial of the matrix A, then $p(A) = A^n + a_{n-1}A^{n-1} + \ldots + a_0 I = 0$. A consequence is that A^n is a linear combination of $\{A^j, \;\; j = 0, \ldots, n-1\}$ and hence A^k, $k \geq n$ is also. Since

$$e^{At} = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!}$$

the Cayley-Hamilton theorem allows us to conclude that

$$e^{At} = \varphi_0(t)I + \varphi_1(t)A + \ldots + \varphi_{n-1}(t)A^{n-1}$$

for certain functions $\{\varphi_i(t)\}$. Let $x \in \mathcal{R}(L_c)$. Then there exists a control u such that

$$\begin{aligned}
x &= \int_0^t e^{A(t-\tau)} Bu(\tau) d\tau \\
&= \int_0^t [\varphi_0(t-\tau)I + \ldots + \varphi_{n-1}(t-\tau)A^{n-1}] Bu(\tau) d\tau .
\end{aligned}$$

Let

$$v_j = \int_0^t \varphi_j(t-\tau)u(\tau)d\tau \,.$$

Then

$$x = [B \;\; AB \; \dots \; A^{n-1}B]\begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{n-1} \end{bmatrix} \in \mathcal{R}(Q_c)\,.$$

Second, we show that $\mathcal{R}(Q_c) \subset \mathcal{R}(L_c)$. Using facts about adjoints and orthogonal subspaces found in the Appendix, we have that

$$\begin{aligned} \mathcal{R}(L_c) &= \mathcal{R}(L_c L_c^*) = \mathcal{N}(L_c L_c^*)^{\perp} \\ \mathcal{R}(Q_c) &= \mathcal{N}({Q_c}^T)^{\perp}\,. \end{aligned}$$

From these facts, showing that $\mathcal{R}(Q_c) \subset \mathcal{R}(L_c)$ is equivalent to showing that

$$\mathcal{N}(L_c L_c^*) \subset \mathcal{N}(Q_c^T)\,.$$

To this end, let $x \in \mathcal{N}(L_c L_c^*)$. Then we have

$$\begin{aligned} 0 &= x^T W_c x \\ &= \int_0^t x^T e^{A\tau} B B^T e^{A^T\tau} x d\tau \\ &= \int_0^t \|B^T e^{A^T\tau} x\|^2 d\tau\,. \end{aligned}$$

This yields

$$B^T e^{A^T\tau} x = 0\,, \qquad 0 \le \tau \le t\,.$$

Setting $\tau = 0$ gives

$$B^T x = 0$$

For $k = 1, \dots, n-1$, take the kth derivative of $B^T e^{A^T\tau}$ with respect to τ and evaluate the result at $\tau = 0$. This gives successively

$$\begin{aligned} B^T A^T x &= 0 \\ B^T (A^T)^2 x &= 0 \\ &\vdots \\ B^T (A^T)^{n-1} x &= 0 \end{aligned}$$

so that $Q_c^T x = 0$. That is $x \in \mathcal{N}(Q_c^T)$, as desired. □

We summarize our results as follows. We say a linear system $\dot{x} = Ax + Bu$ or the pair (A, B) is controllable if any one (hence all) of the following conditions holds:

$$\begin{array}{rrcl} (i) & \mathcal{R}(L_c) & = & \mathbb{R}^n \quad \text{for some (hence all)} \quad t > 0. \\ (ii) & \mathcal{R}(Q_c) & = & \mathbb{R}^n. \\ (iii) & rank\;\; [B \; AB \; \cdots \; A^{n-1}B] & = & n. \end{array}$$

Note that if the system is controllable, then the state can be transferred from any state x_0 (not just 0) at $\tau = 0$ to any other state x_1 at time $\tau = t_1$. This is because if we want to transfer the state from x_0 to x_1, we can simply use the control which transfers the state from 0 to $z = x_1 - e^{At_1}x_0$. In fact, a control input that achieves the transfer is given by

$$u(\tau) = B^T e^{A^T(t_1-\tau)} W_c^{-1}(t_1)(x_1 - e^{At_1}x_0)\,.$$

You can verify for yourself that this control achieves the transfer.

3.2 Alternate Proof of Controllability

Controllability is a deep property of a control system. The proof in the previous section, while the most common one in the textbooks, has the shortcomings that it requires background on linear operators and adjoints and that it does not give much intuition about controllability. Also, it is not easily extended to the nonlinear setting. In this section, we give another proof that a system is controllable if and only if all states are reachable from the origin, in order to provide an alternate view and further insight.

Theorem 3.2.1. *$\mathcal{R}(L_c) = \mathbb{R}^n$ if and only if $rank(Q_c) = n$.*

Proof. (Necessity) We show that if $\mathcal{R}(L_c) = \mathbb{R}^n$ then $rank(Q_c) = n$. Suppose $rank(Q_c) < n$. Consider the expression

$$e^{-At_1}x_1 - x_0 = \int_0^{t_1} e^{-A\tau}Bu(\tau)d\tau\,.$$

By assumption there exists a non-zero vector $v \in \mathbb{R}^n$ such that $v^TQ_c = 0$. This implies

$$v^TB = 0\,, \quad v^TAB = 0\,, \quad \ldots \qquad , v^TA^{n-1}B = 0\,.$$

By the Cayley-Hamilton theorem (see the Appendix), we obtain $v^TA^kB = 0$, for all $k = 0, 1, \ldots$. It follows that $v^Te^{-A\tau}B = 0$ since $e^{-A\tau}B = B - AB\tau + A^2B\frac{\tau^2}{2!} + \cdots$. Therefore

$$v^T\left(e^{-At_1}x_1 - x_0\right) = \int_0^{t_1} v^Te^{-A\tau}Bu(\tau)d\tau = 0\,.$$

This means there is a constraint on x_0 and x_1. But x_0 and x_1 must be arbitrary because $\mathcal{R}(L_c) = \mathbb{R}^n$. Thus, we arrive at a contradiction.

(Sufficiency) We show that if $rank(Q_c) = n$ then $\mathcal{R}(L_c) = \mathbb{R}^n$. Suppose not. That is, suppose the map L_c is not onto. Equivalently, the linear map

$$e^{-At_1}L_c = \int_0^{t_1} e^{-A\tau}Bu(\tau)d\tau$$

is not onto. This means there is a non-zero vector $v \in \mathbb{R}^n$ such that

$$v^T\int_0^{t_1} e^{-A\tau}Bu(\tau)d\tau = 0\,.$$

Choose a control of the form $u(\tau) = (0, \ldots, 0, u_i^s(\tau), 0, \ldots, 0)$ where the ith component is

$$u_i^s(\tau) = \begin{cases} 1 & 0 \le \tau \le 2 \\ 0 & \tau > s \end{cases}$$

and $s \in \mathbb{R}$ is a parameter. Then we have

$$v^T\int_0^{s} e^{-A\tau}b_i d\tau = 0\,, \qquad i = 1, \ldots, m$$

where b_i is the ith column of B. This expression holds for all $s \in \mathbb{R}$. This means

$$v^Te^{-As}b_i = 0\,, \qquad s \in \mathbb{R}, \qquad i = 1, \ldots, m\,.$$

Differentiating this expression repeatedly with respect to s and setting $s = 0$ we obtain

$$v^TA^kb_i = 0\,, \qquad i = 1, \ldots, m$$

for all $k = 0, 1, \ldots$ Equivalently, $v^TQ_c = 0$, which contradicts $rank(Q_c) = n$. □

3.3 Invariance under Change of Basis

Recall that if x is a state vector, so is $T^{-1}x$ for any nonsingular matrix T. In fact, if we let $z = T^{-1}x$,

$$\begin{aligned} \dot{z} &= T^{-1}\dot{x} \\ &= T^{-1}ATz + T^{-1}Bu \end{aligned}$$

so that with z as the state vector, the system matrices change from (A, B) to $(T^{-1}AT, T^{-1}B)$. We refer to this as a change of basis because if we let the columns of T form a new basis for $\mathbb{R}^n$, z is the representation of x in this new basis.

Theorem 3.3.1. *(A, B) is controllable if and only if $(T^{-1}AT, T^{-1}B)$ is controllable for every nonsingular T.*

Proof. Consider the controllability matrix $\tilde{Q}_c$ for the pair $(T^{-1}AT, T^{-1}B)$:

$$\begin{aligned} \tilde{Q}_c &= [T^{-1}B \;\; T^{-1}ATT^{-1}B \cdots] \\ &= T^{-1}[B \;\; AB \cdots] \\ &= T^{-1}Q_c\,. \end{aligned}$$

Since $rank(T^{-1}Q_c) = rank(Q_c)$ the result is proved. □

If the pair (A, B) is not controllable, there is a particular basis in which the controllable and uncontrollable parts are displayed transparently. We illustrate the choice of basis and the computation involved with the following example.

Example 3.3.1.

For the two cart-one force system, if we take $M_1 = K = 1$, $M_2 = \frac{1}{2}$, we obtain the system matrices

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

We know that this system is not controllable, and that the first two columns of the controllability matrix span $\mathcal{R}(Q_c)$. It is easily verified that we can take the following two vectors as a basis for $\mathcal{R}(Q_c)$

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ -2 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ -2 \end{bmatrix}.$$

We complete this to a basis in $\mathbb{R}^4$ by augmenting with, say, the vectors

$$v_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Define the matrix

$$T = \begin{bmatrix} | & | & | & | \\ v_1 & v_2 & v_3 & v_4 \\ | & | & | & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ 0 & -2 & 0 & 1 \end{bmatrix}.$$

With respect to this basis, the state x is transformed to $z = T^{-1}x$. Since

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

the equation governing z is given by

$$\dot{z} = \tilde{A}z + \tilde{B}u$$

where

$$\begin{aligned} \tilde{A} &= T^{-1}AT \\ &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{bmatrix} \end{aligned} \tag{3.1}$$

where $\tilde{A}_{11}$ is the upper left 2x2 block, $\tilde{A}_{12}$ the upper right 2x2 block, and $\tilde{A}_{22}$ the lower right 2x2 block.

$$\begin{aligned} \tilde{B} &= T^{-1}B \\ &= \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} \tilde{B}_1 \\ 0 \end{bmatrix}. \end{aligned} \tag{3.2}$$

Note that the pair $(\tilde{A}_{11}, \tilde{B}_1)$ is controllable because $rank(Q_c) = rank([\tilde{B}_1 \ \tilde{A}_{11}\tilde{B}_1 \ \tilde{A}_{11}^2\tilde{B}_1])$, while z_3 and z_4 corresponding to the $\tilde{A}_{22}$ block is uncontrollable, since they are decoupled from z_1 and z_2 and are unaffected by u.

In general, whenever Q_c is not full rank, we can find a basis so that in the new basis, A and B take the form given in (3.1) and (3.2), respectively, with $(\tilde{A}_{11}, \tilde{B}_1)$ controllable. The procedure is:

1. Find a basis for $\mathcal{R}(Q_c)$. Denote the vectors in this basis by $\{v_1, v_2, \cdots, v_k\}$.
2. Complete the basis to form a basis for $\mathbb{R}^n$. Define the matrix T to have as its columns the basis vectors $\{v_1, v_2, \cdots, v_n\}$.
3. Compute $\tilde{A} = T^{-1}AT$ and $\tilde{B} = T^{-1}B$. $\tilde{A}$ will take the form (3.1) and $\tilde{B}$ will take the form (3.2).

3.4 Invariance under State Feedback

A control law of the form

$$u(t) = Kx(t) + v(t)$$

with $v(t)$ a new input, is referred to as a *state feedback*. The closed-loop system equation is given by

$$\dot{x} = (A + BK)x(t) + Bv(t)$$

It is an important property that controllability is unaffected by state feedback.

Theorem 3.4.1. *(A, B) is controllable if and only if $(A + BK, B)$ is controllable for all K.*

A proof can be obtained using the PBH test. For details, refer to the problem sets.

3.5 Controllable Canonical Form

For single input systems, there is a special form of system matrices for which controllability always holds. This special form is referred to as the *controllable canonical form*. Using a lower case b to indicate explicitly that the input matrix is a column vector for a single input system, the controllable canonical form is given by

$$A = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & \cdots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & \cdots & & -\alpha_{n-1} \end{bmatrix}$$

$$b = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

It is easy to verify that the controllability matrix for this pair (A, b) always has rank n, regardless of the values of the coefficients α_j; hence the name controllable canonical form. An A matrix taking the above form is referred to as a *companion form* matrix. It is straightforward to show that the characteristic polynomial of the companion form matrix is given by

$$\det(sI - A) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_0\,.$$

It will be seen in the next chapter that when using pole assignment in single-input systems, the controllable canonical form is particularly convenient for control design. To prepare for that discussion, we have that

Theorem 3.5.1. *If (A, b) is controllable there exists a similarity transformation T such that $(T^{-1}AT, T^{-1}b)$ is in controllable canonical form.*

Proof. Consider the matrix

$$\begin{aligned}
T &= [A^{n-1}b \;\; A^{n-2}b \;\; \cdots \;\; b] \begin{bmatrix} 1 & 0 & \dots & & 0 \\ \alpha_{n-1} & & & & \\ \alpha_{n-2} & & & & \vdots \\ \vdots & & & & \\ \alpha_2 & & & & 0 \\ \alpha_1 & \alpha_2 & \dots & \alpha_{n-1} & 1 \end{bmatrix} \\
&= [b \; \cdots \; A^{n-1}b] \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{n-1} & 1 \\ \alpha_2 & \cdots & \alpha_{n-1} & 1 & 0 \\ \vdots & \iddots & 1 & 0 & 0 \\ \alpha_{n-1} & \iddots & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}.
\end{aligned}$$

The second matrix (called a Toeplitz matrix) forming T is non-singular. The first matrix forming T is Q_c. Thus, controllability ensures T^{-1} exists, so that its columns $v_1, ..., v_n$ form a basis of $\mathbb{R}^n$.

Note that

$$\begin{aligned}
v_1 &= A^{n-1}b + \alpha_{n-1}A^{n-2}b + ... + \alpha_1 b \\
v_2 &= A^{n-2}b + \alpha_{n-1}A^{n-3}b + \ldots + \alpha_2 b \\
&\vdots \\
v_{n-1} &= Ab + \alpha_{n-1}b \\
v_n &= b
\end{aligned}$$

and that

$$\begin{aligned}
Av_1 &= A^n b + ... + \alpha_1 Ab + \alpha_0 b - \alpha_0 b \\
&= -\alpha_0 b \qquad \text{by the Cayley-Hamilton Theorem} \\
&= -\alpha_0 v_n \\
Av_2 &= v_1 - \alpha_1 v_n \\
&\vdots \\
Av_n &= v_{n-1} - \alpha_{n-1} v_n
\end{aligned}$$

Thus the matrix representation of A with respect to the basis $\{v_1 ... v_n\}$ looks like

$$\tilde{A} = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ \vdots & & & \ddots & \\ 0 & 0 & \cdots & 0 & 1 \\ -\alpha_0 & -\alpha_1 & \cdots & & -\alpha_{n-1} \end{bmatrix}$$

Similarly, the vector b looks like

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} = \tilde{b}$$

But $\tilde{A}$ and $\tilde{b}$ are then related to the original matrices through

$$\begin{aligned} \tilde{A} &= T^{-1}AT \\ \tilde{b} &= T^{-1}b \end{aligned}$$

so that they are related by a similarity transformation. Thus the new system $z(t) = T^{-1}x(t)$ will satisfy an equation of the form

$$\dot{z} = \tilde{A}z + \tilde{b}u$$

with $(\tilde{A}, \tilde{b})$ in controllable canonical form. □

3.6 PBH Test

There is a very useful test for controllability, referred to as the PBH test.

Theorem 3.6.1 (PBH). *(A, B) is controllable if and only if rank$[A - \lambda I \quad B] = n$ for all eigenvalues λ of A.*

This theorem can be proved using the change of basis described above for uncontrollable systems. Details are provided in the problem sets. It is important to note that $Rank[A-\lambda I \quad B] = n$ for all eigenvalues λ of A if and only if $Rank[A - \lambda I \quad B] = n$ for all complex numbers λ. This is because for λ not an eigenvalue of A, $Rank(A - \lambda I) = n$.