Control Theory

Course ID: 322MI – Spring 2023

Felice Andrea Pellegrino

University of Trieste
Department of Engineering and Architecture

322MI –Spring 2023

Lecture 4: Structural properties and special forms

Reachability and controllability for continuous-time linear time-invariant systems

Definitions

It is frequently desirable to **find an input that causes the state of a system to assume specified values in finite time** (for instance, transfer the state vector from $\bar{x}_1$ to $\bar{x}_2$). This type of desirable property leads to the concepts of reachability and controllability. Indeed, reachability and controllability concern how the input affects the state of a system.

The following definitions concern systems of the form:

$$\dot{x}(t) = Ax(t) + Bu(t)$$

(since we are dealing with properties concerning input and state, the output transform can be ignored).

Definitions (cont.)

Definition: reachability and controllability on an interval

Given the states $\bar{x}_1$ and $\bar{x}_2$, and the interval $[t_1, t_2]$, if an input $u : [t_1, t_2] \to \mathbb{R}^m$ exists such that

$$x(t_1) = \bar{x}_1 \quad \Longrightarrow \quad x(t_2) = \bar{x}_2$$

then we say that $\bar{x}_2$ is *reachable from* $\bar{x}_1$ on the interval $[t_1, t_2]$ and $\bar{x}_1$ is *controllable to* $\bar{x}_2$ on the same interval.

It is often useful to take $\bar{x}_1 = 0$ and $t_1 = 0$, thus we have the following.

Definition: reachability from 0

The state vector $\bar{x} \in \mathbb{R}^n$, is said to be *reachable* (from 0) on the interval $[0, \tau]$ if $\exists u : [0, \tau] \to \mathbb{R}^m$ such that

$$x(0) = 0 \quad \implies \quad x(\tau) = \bar{x}$$

Definition: controllability to 0

The state vector $\bar{x} \in \mathbb{R}^n$, is said to be *controllable* (to 0) on the interval $[0, \tau]$ if $\exists u : [0, \tau] \to \mathbb{R}^m$ such that

$$x(0) = \bar{x} \quad \implies \quad x(\tau) = 0$$

Reachable set and controllable set

Definition: reachable set

Given the time τ, the reachable set on $[0, \tau]$, denoted by $X_r(\tau)$, consists of all states $\bar{x}$ for which there exists an input $u : [0, \tau] \to \mathbb{R}^m$ that transfers the state from $x(0) = 0$ to $x(\tau) = \bar{x}$, i.e.,

$$X_r(\tau) = \{\bar{x} \in \mathbb{R}^n : \bar{x} \text{ is reachable from 0 in } [0, \tau]\}$$

Definition: controllable set

Given the time τ, the controllable (to 0) set on $[0, \tau]$, denoted by $X_c(\tau)$, consists of all states $\bar{x}$ for which there exists an input $u : [0, \tau] \to \mathbb{R}^m$ that transfers the state from $x(0) = \bar{x}$ to $x(\tau) = 0$, i.e.,

$$X_c(\tau) = \{\bar{x} \in \mathbb{R}^n : \bar{x} \text{ is controllable to 0 in } [0, \tau]\}$$

Reachable set and controllable set (cont.)

Theorem

$X_r(\tau)$ and $X_c(\tau)$ are subspaces of $\mathbb{R}^n \; \forall \, \tau$.

Proof.
We need to prove that if $\bar{x}_1, \bar{x}_2 \in X_r(\tau)$ then $\alpha\bar{x}_1 + \beta\bar{x}_2 \in X_r(\tau), \forall \alpha, \beta$. Since $\bar{x}_1 \in X_r(\tau)$, there exists $u_1 : [0, \tau] \to \mathbb{R}^m$ such that

$$\bar{x}_1 = 0 + \int_0^{\tau} e^{A(t-\sigma)} B u_1(\sigma) d\sigma$$

Similarly, there exists $u_2 : [0, \tau] \to \mathbb{R}^m$ such that

$$\bar{x}_2 = 0 + \int_0^{\tau} e^{A(t-\sigma)} B u_2(\sigma) d\sigma$$

By linearly combining the last equations we get, for any α, β

$$\alpha\bar{x}_1 + \beta\bar{x}_2 = \int_0^{\tau} e^{A(t-\sigma)} B \left[\alpha u_1(\sigma) + \beta u_2(\sigma)\right] d\sigma$$

Hence, $\alpha\bar{x}_1 + \beta\bar{x}_2$ is reachable from 0 (with input $\alpha u_1 + \beta u_2$).

The proof for $X_c(\tau)$ is analogous. □

Characterization of the reachable subspace

Definition: reachability matrix

Given a system

$$\dot{x}(t) = Ax(t) + Bu(t),$$

where $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$, the $n \times (nm)$ matrix

$$R = [B \mid AB \mid A^2B \mid \dots \mid A^{n-1}B]$$

is said the *reachability matrix*.

Theorem

For all $\tau > 0$, the reachable subspace $X_r(\tau)$ is independent of τ and is the image of the reachability matrix:

$$X_r(\tau) = X_r = \operatorname{im} R.$$

Proof.

The reachable subspace on the interval $[0, \tau]$ is

$$X_r(\tau) = \left\{ x : x = \int_0^\tau e^{A(\tau-\sigma)} B u(\sigma) d\sigma, \ \text{ for some } u : [0, \tau] \to \mathbb{R}^m \right\} \tag{1}$$

To prove that $X_r(\tau) = \operatorname{im} R$, we will show that the orthogonal complements are equal:

$$X_r^\perp(\tau) = (\operatorname{im} R)^\perp .$$

The orthogonal set to $X_r(\tau)$ is $X_r^\perp(\tau) = \{z : z^\top x = 0 \ \forall x \in X_r(\tau)\}$ and the condition $z^\top x = 0 \ \forall x \in X_r(\tau)$ implies

$$z^\top \int_0^\tau e^{A(\tau-\sigma)} B u(\sigma) d\sigma = \int_0^\tau z^\top e^{A(\tau-\sigma)} B u(\sigma) d\sigma = 0, \quad \forall u : [0, \tau] \to \mathbb{R}^m$$

Since $u(\sigma)$ is arbitrary, it follows that $z^\top e^{A(\tau-\sigma)}B$ is identically zero for any σ in $[0,\tau]$, namely

$$z^\top e^{A(\tau-\sigma)}B \equiv 0$$

(otherwise, we could take $u(\sigma) = (z^\top e^{A(\tau-\sigma)}B)^\top \doteq W^\top(\sigma)$ thus obtaining $\int_0^\tau W(\sigma)W^\top(\sigma)d\sigma = \int_0^\tau \|W(\sigma)\|^2\, d\sigma > 0$).

Considering the expression of the matrix exponential, we can write:

$$\begin{aligned} z^\top e^{A(\tau-\sigma)}B &= z^\top \left[I + A(\tau-\sigma) + \frac{[A(\tau-\sigma)]^2}{2!} + \dots \right] B \\ &= \sum_{k=0}^{\infty} z^\top \frac{[A(\tau-\sigma)]^k}{k!} B = 0, \qquad \forall \sigma \in [0,\tau] \end{aligned}$$

By the **principle of identity for power series** this equality is true if and only if

$$z^\top A^k B = 0 \quad \forall k \geq 0 \,.$$

The latter condition can be shown to be equivalent to

$$z^\top A^k B = 0 \quad \forall k \in [0, n-1] \,.$$

Characterization of the reachable subspace (cont.)

The direct implication is trivial. The opposite is a consequence of the Cayley-Hamilton theorem, that implies that the powers A^k with $k \geq n$ can be expressed as a linear combination of the first n powers of A i.e., I, A, A^2, ..., A^{n-1}. Therefore, $z^\top A^k B$ is a linear combination of the terms $z^\top B$, $z^\top AB$, $z^\top A^2 B$, ..., $z^\top A^{n-1} B$.

Thus, the orthogonality condition becomes:

$$z^\top \underbrace{[B \mid AB \mid A^2 B \mid \ \ldots \ \mid A^{n-1} B]}_{R} = 0 \quad \Longleftrightarrow \quad z^\top R = 0\,.$$

In turn, we have proven that:

$$X_r^\perp(\tau) = \{z : z^\top R = 0\}$$

Hence, $X_r^\perp(\tau)$ is the set of all the vectors orthogonal to the columns of R, i.e., orthogonal to $\operatorname{im} R$. In other words:

$$X_r^\perp(\tau) = (\operatorname{im} R)^\perp$$

which is equivalent to

$$X_r(\tau) = \operatorname{im} R.$$

□

Observation

The fact that the reachability does not depend on the time horizon deserves some discussion. Indeed, it implies that any reachable state can be reached in an arbitrarily small amount of time. This is clearly not true for many physical systems. For example, a car cannot reach any state in an arbitrarily small amount of time. The apparent contradiction can be explained by the fact that the theorem does not take into account upper bounds on the magnitude of the input. If such bounds are taken into account (for instance, $\|u\| \leq u_{\text{max}}$), it can be shown that $X_r(\tau)$ is a bounded set that depends on τ.

An algebraic condition for reachability

The question arises whether, given a system, all its states are reachable.

Definition: reachable system

A system is said to be *reachable* if all its states are reachable.

Since $X_r = \operatorname{im} R$, we have the following fundamental result:

Theorem

A system is reachable if and only if

$$\operatorname{rank} R = n.$$

Characterization of the controllable subspace

Theorem

For all $\tau > 0$, the controllable subspace $X_c(\tau)$ is independent of τ and is the image of the reachability matrix:

$$X_c(\tau) = X_c = \operatorname{im} R.$$

Proof.

The controllable subspace on the interval $[0, \tau]$ is:

$$X_c(\tau) = \left\{ x : \; e^{A\tau} x + \int_0^\tau e^{A(\tau-\sigma)} B u(\sigma) d\sigma = 0, \quad \text{for some } u : [0, \tau] \to \mathbb{R}^m \right\}.$$

Since the matrix $e^{A\tau}$ is invertible (its inverse being $e^{-A\tau}$), we can solve for x:

$$x = -e^{-A\tau} \int_0^\tau e^{A(\tau-\sigma)} B u(\sigma) d\sigma = -\int_0^\tau e^{-A\sigma} B u(\sigma) d\sigma,$$

where the last equality follows from the fact that $e^{A\tau}$ and $e^{-A\sigma}$ commute. Thus, $X_c(\tau)$ can be characterized as:

$$X_c(\tau) = \left\{ x : x = -\int_0^\tau e^{-A\sigma} B u(\sigma) d\sigma, \quad \text{for some } u : [0, \tau] \to \mathbb{R}^m \right\}$$

which is the same as $\operatorname{im} R$ (the rest of the proof is similar to the one for $X_r(\tau)$). □

An algebraic condition for controllability

The question arises whether, given a system, all its states are controllable.

Definition: controllable system

A system is said to be *controllable* if all its states are controllable.

Since $X_c = \operatorname{im} R$, we have the following fundamental result:

Theorem

A system is controllable if and only if

$$\operatorname{rank} R = n.$$

Kalman decomposition for reachability

Suppose a system is not reachable and let X_r be the reachable subspace, of dimension $r < n$.

Take a basis $t_1, \dots, t_r$ for X_r, i.e.,

$$X_r = < t_1, \dots, t_r >$$

(for instance, take r linearly independent columns of R) and form an $n \times r$ matrix

$$T_r = [t_1 \ \dots \ t_r].$$

Then, complete (arbitrarily) a basis of $\mathbb{R}^n$ by choosing $n - r$ vectors such that

$$\mathbb{R}^n = < t_1, \dots, t_r, \underbrace{t_{r+1}, \dots, t_n}_{\text{arbitrary completion}} >,$$

and form the matrix

$$T_{\bar{r}} = [t_{r+1} \ \dots \ t_n].$$

Each vector $x \in \mathbb{R}^n$ can thus be expressed as a linear combination of the columns of T_r and $T_{\bar{r}}$ as follows:

$$x = T_r \hat{x}_r + T_{\bar{r}} \hat{x}_{\bar{r}} \ ,$$

Kalman decomposition for reachability (cont.)

where $\hat{x}_r \in \mathbb{R}^r$ and $\hat{x}_{\bar{r}} \in \mathbb{R}^{n-r}$. More compactly:

$$x = [T_r \mid T_{\bar{r}}] \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} \tag{2}$$

In particular, each vector belonging to the reachable subspace can be written as:

$$x = [T_r \mid T_{\bar{r}}] \begin{bmatrix} \hat{x}_r \\ 0 \end{bmatrix}$$

If we apply the state transformation $T = [T_r \mid T_{\bar{r}}]$ to the system, we get the new representation:

$$\begin{cases} \dot{\hat{x}}(t) & = & \hat{A}\hat{x}(t) + \hat{B}u(t) \\ y(t) & = & \hat{C}\hat{x}(t) + \hat{D}u(t) \end{cases}$$

where $\hat{A} = T^{-1}AT, \hat{B} = T^{-1}B, \hat{C} = CT, \hat{D} = D$.

Kalman decomposition for reachability (cont.)

The new state vector is thus

$$\hat{x} = [T_r \mid T_{\bar{r}}]^{-1} x$$

where, from (2):

$$\hat{x} = \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} .$$

In the new representation, a vector $\hat{x}$ is reachable if and only if $\hat{x}_{\bar{r}} = 0$. Indeed, the two components $\hat{x}_r$ and $\hat{x}_{\bar{r}}$ represent the "reachable part" and the "non-reachable part", respectively.

The state equation can be written as

$$\begin{bmatrix} \dot{\hat{x}}_r \\ \dot{\hat{x}}_{\bar{r}} \end{bmatrix} = \begin{bmatrix} A_r & A_{r,\bar{r}} \\ \phi_1 & A_{\bar{r}} \end{bmatrix} \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} + \begin{bmatrix} B_r \\ \phi_2 \end{bmatrix} u . \qquad (3)$$

We demonstrate now that ϕ_1 and ϕ_2 are null matrices.

Kalman decomposition for reachability (cont.)

By assuming zero initial conditions

$$\begin{bmatrix} \hat{x}_r(0) \\ \hat{x}_{\bar{r}}(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

the state at any $t \geq 0$ must belong to the reachable subspace, thus be of the form:

$$\begin{bmatrix} \hat{x}_r \\ 0 \end{bmatrix}.$$

In particular

$$\hat{x}_{\bar{r}}(t) = 0, \quad \forall t \geq 0 \qquad \Longrightarrow \qquad \dot{\hat{x}}_{\bar{r}}(t) = 0, \quad \forall t \geq 0$$

Thus, the second equation of (3)

$$\dot{\hat{x}}_{\bar{r}}(t) = \phi_1 \hat{x}_r(t) + A_{\bar{r}} \hat{x}_{\bar{r}}(t) + \phi_2 u(t),$$

can be written, $\forall\, t \geq 0$, as

$$0 = \phi_1 \hat{x}_r(t) + \phi_2 u(t) = \begin{bmatrix} \phi_1 & \phi_2 \end{bmatrix} \begin{bmatrix} \hat{x}_r(t) \\ u(t) \end{bmatrix}.$$

Kalman decomposition for reachability (cont.)

Notice that

- $\hat{x}_r(t)$, being reachable, can be arbitrarily chosen by a proper choice of the input up to time t;
- $u(t)$ can be arbitrarily chosen.

The only matrix which gives zero when multiplied by any vector is the null matrix, hence

$$\begin{bmatrix} \phi_1 & \phi_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \end{bmatrix}.$$

Therefore, the system can be written in the form

$$\begin{aligned} \begin{bmatrix} \dot{\hat{x}}_r \\ \dot{\hat{x}}_{\bar{r}} \end{bmatrix} &= \begin{bmatrix} A_r & A_{r,\bar{r}} \\ 0 & A_{\bar{r}} \end{bmatrix} \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} + \begin{bmatrix} B_r \\ 0 \end{bmatrix} u \\ y &= \begin{bmatrix} C_r & C_{\bar{r}} \end{bmatrix} \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} + Du \end{aligned} \tag{4}$$

(the blocks C_r and $C_{\bar{r}}$ do not have any particular property).

We form just obtained:

$$\begin{bmatrix} \dot{\hat{x}}_r \\ \dot{\hat{x}}_{\bar{r}} \end{bmatrix} = \begin{bmatrix} A_r & A_{r,\bar{r}} \\ 0 & A_{\bar{r}} \end{bmatrix} \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} + \begin{bmatrix} B_r \\ 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} C_r & C_{\bar{r}} \end{bmatrix} \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} + Du$$

deserves some analysis. Indeed, we recognize a reachable (see next page) subsystem

$$\Sigma(A_r, B_r, C_r),$$

and a non-reachable subsystem

$$\Sigma(A_{\bar{r}}, 0, C_{\bar{r}}).$$

The latter is clearly non-reachable, since its input matrix is null. The former will be proven to be reachable next. Finally, the block $A_{r,\bar{r}}$ represents the action of the non-reachable subsystem on reachable one.

Kalman decomposition for reachability (cont.)

To prove that

$$\Sigma(A_r, B_r, C_r)$$

is reachable, we can compute the reachability matrix of the whole system in the new basis:

$$\hat{R} = \begin{bmatrix} \hat{B} & \hat{A}\hat{B} & \cdots & \hat{A}^{n-1}\hat{B} \end{bmatrix} = \\ = \left[\begin{array}{c|c|c|c} \hat{B}_r & \hat{A}_r \hat{B}_r & & \hat{A}_r^{n-1} \hat{B}_r \\ & & \cdots & \\ 0 & 0 & & 0 \end{array}\right]$$

Observing that

$$\operatorname{rank} \hat{R} = \operatorname{rank} \left(T^{-1} R\right) = \operatorname{rank} R = r,$$

we have:

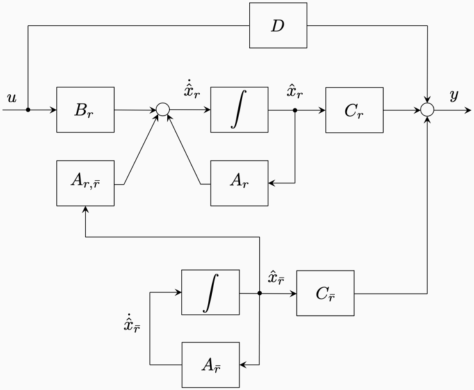
$$r = \operatorname{rank} \hat{R} \; = \; \operatorname{rank} \begin{bmatrix} \hat{B}_r & \hat{A}_r \hat{B}_r & \cdots & \cdots & \hat{A}_r^{n-1} \hat{B}_r \end{bmatrix} = \\ = \; \operatorname{rank} \underbrace{\begin{bmatrix} \hat{B}_r & \hat{A}_r \hat{B}_r & \cdots & \hat{A}_r^{r-1} \hat{B}_r \end{bmatrix}}_{\hat{R}_r}$$

where the last equality is due to the Cayley-Hamilton theorem. Thus, $\hat{R}_r$ is full rank, meaning that the considered subsystem is reachable.

Kalman decomposition for reachability (cont.)

$$\begin{bmatrix} \dot{\hat{x}}_r \\ \dot{\hat{x}}_{\bar{r}} \end{bmatrix} = \begin{bmatrix} A_r & A_{r,\bar{r}} \\ 0 & A_{\bar{r}} \end{bmatrix} \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} + \begin{bmatrix} B_r \\ 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} C_r & C_{\bar{r}} \end{bmatrix} \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} + Du$$



$$
\begin{bmatrix} \dot{\hat{x}}_r \\ \dot{\hat{x}}_{\bar{r}} \end{bmatrix} = \begin{bmatrix} A_r & A_{r,\bar{r}} \\ 0 & A_{\bar{r}} \end{bmatrix} \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} + \begin{bmatrix} B_r \\ 0 \end{bmatrix} u
$$

$$
y = \begin{bmatrix} C_r & C_{\bar{r}} \end{bmatrix} \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} + Du
$$

- the input u does not affect the component $\hat{x}_{\bar{r}}$ at all, therefore $\hat{x}_{\bar{r}}(t)$ is determined only by the initial value $\hat{x}_{\bar{r}}(0)$:

$$
\hat{x}_{\bar{r}}(t) = e^{A_{\bar{r}} t} \hat{x}_{\bar{r}}(0)
$$

Observations (cont.)

$$\begin{bmatrix} \dot{\hat{x}}_r \\ \dot{\hat{x}}_{\bar{r}} \end{bmatrix} = \begin{bmatrix} A_r & A_{r,\bar{r}} \\ 0 & A_{\bar{r}} \end{bmatrix} \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} + \begin{bmatrix} B_r \\ 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} C_r & C_{\bar{r}} \end{bmatrix} \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} + Du$$

- the set $\sigma(A)$ of the eigenvalues of A (which, by similarity, are the same of $\hat{A}$) can be partitioned into two subsets:

  $$\sigma(A) = \sigma(A_r) \cup \sigma(A_{\bar{r}}),$$

  where

  - $\sigma(A_r) = \{\lambda_1, \dots, \lambda_r\}$ are the *reachable eigenvalues*, associated with *reachable modes*, and
  - $\sigma(A_{\bar{r}}) = \{\lambda_{r+1}, \dots, \lambda_n\}$ are the *unreachable eigenvalues*, associated with *unreachable modes*.

$$\begin{bmatrix} \dot{\hat{x}}_r \\ \dot{\hat{x}}_{\bar{r}} \end{bmatrix} = \begin{bmatrix} A_r & A_{r,\bar{r}} \\ 0 & A_{\bar{r}} \end{bmatrix} \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} + \begin{bmatrix} B_r \\ 0 \end{bmatrix} u$$

$$y = \begin{bmatrix} C_r & C_{\bar{r}} \end{bmatrix} \begin{bmatrix} \hat{x}_r \\ \hat{x}_{\bar{r}} \end{bmatrix} + Du$$

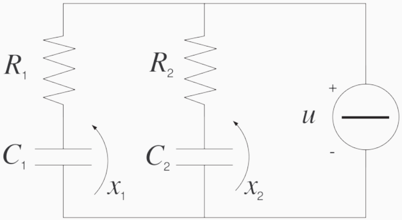
- by recalling that the transfer function is invariant for change of basis, and applying the formula for the inverse of a triangular block matrix:

$$\begin{bmatrix} P & Q \\ 0 & R \end{bmatrix}^{-1} = \begin{bmatrix} P^{-1} & -P^{-1}QR^{-1} \\ 0 & R^{-1} \end{bmatrix}$$

it can be shown that the transfer function depends on the reachable subsystem only

$$W(s) = C(sI - A)^{-1}B + D = C_r(sI - A_r)^{-1}B_r + D.$$

Thus, the poles of $W(s)$ are necessarily eigenvalues of A_r.



For the circuit in figure, take the voltage $u(t)$ as input, and x_1 and x_2 (voltage of capacitors) as state variables.

The current through R_1 is $i_1(t) = \dfrac{u(t) - x_1(t)}{R_1}$. The voltage of C_1 is $x_1(t) = \dfrac{q_1(t)}{C_1}$, thus $\dot{x}_1(t) = \dfrac{\dot{q}_1(t)}{C_1} = \dfrac{i_1(t)}{C_1}$. By substituting, we get

$$\dot{x}_1(t) = -\frac{1}{R_1 C_1} x_1(t) + \frac{1}{R_1 C_1} u(t)$$

and, similarly:

$$\dot{x}_2(t) = -\frac{1}{R_2 C_2} x_2(t) + \frac{1}{R_2 C_2} u(t).$$

The state space representation is

$$\dot{x}(t) = \begin{bmatrix} -\frac{1}{R_1C_1} & 0 \\ 0 & -\frac{1}{R_2C_2} \end{bmatrix} x(t) + \begin{bmatrix} \frac{1}{R_1C_1} \\ \frac{1}{R_2C_2} \end{bmatrix} u(t), \qquad (5)$$

and the reachability matrix is

$$R = [B \ AB] = \begin{bmatrix} \frac{1}{R_1C_1} & -\frac{1}{R_1^2C_1^2} \\ \frac{1}{R_2C_2} & -\frac{1}{R_2^2C_2^2} \end{bmatrix}.$$

By computing $\det(R)$, it is easy to verify that the system is reachable if and only if $R_1C_1 \neq R_2C_2$ (we assume $R_1, R_2, C_1, C_2 > 0$). When $R_1C_1 = R_2C_2$, the rank of R is 1, thus the system is not reachable. In that case, X_r is a subspace of dimension 1, given by

$$x = \alpha \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \alpha \in \mathbb{R}.$$

The vector $t_1 = [1 \ \ 1]^\top$ is a basis for X_r. To get the Kalman form it is sufficient to complete the basis, for instance by taking $t_2 = [-1 \ \ 1]^\top$, and perform a change of basis using $T = [t_1 \ t_2]$.

Popov-Belevitch-Hautus (PBH) test for reachability

The Kalman decomposition allows to easily prove the following very useful result, known as *Popov-Belevitch-Hautus (PBH) test for reachability*.

Theorem

The system

$$\dot{x}(t) \quad = \quad Ax(t) + Bu(t)$$

is reachable if and only if

$$\text{rank}\,[\lambda I - A \mid B] = n, \quad \forall \lambda \in \mathbb{C}.$$

Proof.
(⇒) Suppose by contradiction that the system is reachable and $\text{rank}[\lambda I - A \mid B] < n$ for some $\lambda \in \mathbb{C}$. Then, a vector $z \neq 0$ exists such that $z^\top[\lambda I - A \mid B] = 0$. As a consequence we have $z^\top B = 0$ and $z^\top[\lambda I - A] = 0$ (which can be rewritten as $z^\top \lambda = z^\top A$).

Thus, we can write

$$z^\top AB = \lambda z^\top B = 0.$$

By induction, assume $z^\top A^k B = 0$; it follows that

$$z^\top A^{k+1} B = z^\top A A^k B = \lambda z^\top A^k B = 0.$$

Thus, $z^\top A^k B = 0$ for all $k > 0$, and in particular for $k = 0, 1, \dots, n-1$:

$$[z^\top B \mid z^\top AB \mid z^\top A^2 B \mid \dots \mid z^\top A^{n-1} B] = z^\top [B \mid AB \mid \dots | A^{n-1} B] = z^\top R = 0,$$

which implies that the system is not reachable.

($\Leftarrow$) Now we show that, if the system is not reachable, there exists $\lambda \in \mathbb{C}$ such that $\mathrm{rank}[\lambda I - A \mid B] < n$.

Let T be the change of basis matrix that leads to the Kalman reachability form:

$$T^{-1}AT = \begin{bmatrix} A_r & A_{r,\bar{r}} \\ 0 & A_{\bar{r}} \end{bmatrix}, \qquad T^{-1}B = \begin{bmatrix} B_r \\ 0 \end{bmatrix}.$$

We can write:

$$\begin{aligned} \mathrm{rank}\,[\lambda I - A \mid B] &= \mathrm{rank}\; T^{-1}[\lambda I - A \mid B] \begin{bmatrix} T & 0 \\ 0 & I \end{bmatrix} \\ &= \mathrm{rank}\,\left[\lambda T^{-1}T - T^{-1}AT \;\mid\; T^{-1}B\right] \\ &= \mathrm{rank} \left[\begin{array}{cc|c} \lambda I_{n_1} - A_r & -A_{r,\bar{r}} & B_r \\ 0 & \lambda I_{n_2} - A_{\bar{r}} & 0 \end{array}\right] \end{aligned}$$

where n_2 is the dimension of the unreachable subspace $X_{\bar{r}}$. Then, it is sufficient to take $\lambda \in \sigma(A_{\bar{r}})$ to get a rank drop.

□

Popov-Belevitch-Hautus (PBH) test for reachability (cont.)

Observe that:

- it is sufficient to check for $\lambda \in \sigma(A)$, i.e., the set of eigenvalues of A;
- the matrix $[\lambda I - A \mid B]$ has a rank drop for each $\lambda \in \sigma(A_{\bar{r}})$ i.e., for each non-reachable eigenvalue.

Definition

Given the system $\dot{x}(t) = Ax(t) + Bu(t)$, and a time $T > 0$, the matrix

$$W_r(T) = \int_0^T e^{A(T-\tau)} B B^\top e^{A^\top (T-\tau)} d\tau$$

is said to be the *reachability Gramian*.

Clearly, the reachability Gramian is a symmetric positive semi-definite $n \times n$ matrix, for all $T > 0$.

The following result establishes a correspondence between the reachability matrix R and the reachability Gramian $W_r(T)$. It also provides a formula for the control $u(\cdot)$ that reaches a given (reachable) state.

Reachability Gramian (cont.)

Theorem

The image of the reachability Gramian $W_r(T)$ is independent of $T > 0$ and is equal to the image of the reachability matrix R:

$$\operatorname{im} W_r(T) = \operatorname{im} R, \quad \forall T > 0.$$

Moreover, if $\bar{x} = W_r(T)\eta \in \operatorname{im} W_r(T)$, the control

$$u(t) = B^\top e^{A^\top (T-t)}\eta, \quad t \in [0,\ T] \tag{6}$$

steers the state from $x(0) = 0$ to $x(T) = \bar{x}$.

Proof.

We first show that

$$\operatorname{im} W_r(T) \subset \operatorname{im} R.$$

For each $\bar{x} \in \operatorname{im} W_r(T)$ there exists η such that

$$\bar{x} = W_r(T)\eta.$$

Now, apply the input (6) starting from $x(0) = 0$.

We get:

$$\begin{aligned} x(T) &= \int_0^T e^{A(T-\tau)} B u(\tau) d\tau \\ &= \int_0^T e^{A(T-\tau)} B B^\top e^{A^\top (T-\tau)} \eta d\tau \\ &= \underbrace{\left(\int_0^T e^{A(T-\tau)} B B^\top e^{A^\top (T-\tau)} d\tau \right)}_{W_r(T)} \eta = \bar{x}. \end{aligned}$$

Thus, any $\bar{x} \in \operatorname{im} W_r(T)$ is reachable, hence it belongs to $\operatorname{im} R$.

Now we show that

$$\operatorname{im} R \subset \operatorname{im} W_r(T).$$

We need to prove that

$$\bar{x} \in \operatorname{im} R \quad \Longrightarrow \quad \bar{x} \in \operatorname{im} W_r(T) = (\ker W_r(T))^{\perp},$$

where the last equality follows from the identity

$$\operatorname{im} M = \left(\ker M^{\top}\right)^{\perp}$$

and the fact that the Gramian is symmetric.

If $\bar{x} \in \operatorname{im} R$, there exists an input $u(\cdot)$ for which

$$\bar{x} = \int_0^T e^{A(T-\tau)} B u(\tau) d\tau.$$

Take an arbitrary vector $\eta \in \ker W_r(T) = \ker W_r^{\top}(T)$ and compute

$$\bar{x}^{\top}\eta = \int_0^T u^{\top}(\tau) B^{\top} e^{A^{\top}(T-\tau)} \eta d\tau. \tag{7}$$

On the other hand, since $\eta \in \ker W_r(T) \Rightarrow W_r(T)\eta = 0$, we can write

$$\eta^\top W_r(T)\eta = \int_0^T \eta^\top e^{A(T-\tau)} BB^\top e^{A^\top (T-\tau)} \eta d\tau$$
$$= \int_0^T \|B^\top e^{A^\top (T-\tau)} \eta\|^2 d\tau = 0.$$

As a consequence:

$$B^\top e^{A^\top (T-\tau)} \eta = 0, \quad \forall \tau \in [0, \ T].$$

From the previous and (7) we conclude that

$$\bar{x}^\top \eta = 0, \quad \forall \eta \in \ker W_r(T).$$

Hence, any $\bar{x} \in \operatorname{im} R$ is orthogonal to $\ker W_r(T) = \ker W_r^\top(T)$, and thus it belongs to $\operatorname{im} W_r(T)$. □

Reachability Gramian (cont.)

The following corollary provides a condition which is equivalent to reachability.

Corollary

The system $\dot{x}(t) = Ax(t) + Bu(t)$ is reachable if and only if

$$\operatorname{rank} W_r(T) = n \quad \forall T > 0,$$

i.e., the reachability Gramian is nonsingular for all $T > 0$.

The following corollary (whose proof is left as an exercise) establishes the importance of reachability in determining an input u to transfer the state from any x_0 to any $\bar{x}$ in finite time.

Corollary

Let the system $\dot{x}(t) = Ax(t) + Bu(t)$ be reachable, or the pair (A, B) be reachable. Then there exist an input that will transfer any state x_0 to any other state $\bar{x}$ in some finite time T. Such input is given by:

$$u(t) = B^\top e^{A^\top (T-t)} W_r^{-1}(T)[\bar{x} - e^{AT} x_0], \quad t \in [0,\ T].$$

Example

The system $\dot{x}(t) = Ax(t) + Bu(t)$, with $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is reachable. A control input that will transfer any state x_0 to any other state $\bar{x}$ in the finite time T is given by:

$$\begin{aligned} u(t) &= B^\top e^{A^\top (T-t)} W_r^{-1}(T)[\bar{x} - e^{AT} x_0] \\ &= [T - t \quad 1] \begin{bmatrix} 12/T^3 & -6/T^2 \\ -6/T^2 & 4/T \end{bmatrix} \left(\bar{x} - \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} x_0 \right). \end{aligned}$$

Let the *energy of a control signal* $u(\cdot) : [0\ ,\ T] \longrightarrow \mathbb{R}^m$ be defined as

$$J(u(\cdot)) = \int_0^T \|u(\tau)\|^2\, d\tau = \int_0^T u^\top(\tau)u(\tau)d\tau. \tag{8}$$

The control law of the previous theorem has the remarkable property of being a **minimum energy control**.

More precisely, if $\bar{x} = W_r(T)\eta$, then the control

$$u(t) = B^\top e^{A^\top(T-t)}\eta, \quad t \in [0,\ T] \tag{9}$$

achieves the minimum value of J among all the controls that steer the state from $x(0) = 0$ to $x(T) = \bar{x}$.

Indeed, the control (9) must satisfy:

$$\bar{x} = \int_0^T e^{A(T-\tau)}Bu(\tau)d\tau. \tag{10}$$

Take an arbitrary input $\tilde{u}(\cdot)$ that steers the state to $\bar{x}$ from zero:

$$\bar{x} = \int_0^T e^{A(T-\tau)}B\tilde{u}(\tau)d\tau. \tag{11}$$

By subtracting we get:

$$\int_0^T e^{A(T-\tau)} B \underbrace{(\tilde{u}(\tau) - u(\tau))}_{\doteq \delta u(\tau)} d\tau = 0. \tag{12}$$

Thus, the energy of the input $\tilde{u}(\cdot)$ is:

$$\begin{aligned} J(\tilde{u}(\cdot)) &= \int_0^T \tilde{u}^\top(\tau)\tilde{u}(\tau)d\tau = \int_0^T (u(\tau) + \delta u(\tau))^\top (u(\tau) + \delta u(\tau))d\tau \\ &= \int_0^T u^\top(\tau)u(\tau)d\tau \; + \; 2\underbrace{\int_0^T u^\top(\tau)\delta u(\tau)d\tau}_{= 0 \text{ by (9) and (12)}} \; + \; \int_0^T \underbrace{\delta u^\top(\tau)\delta u(\tau)}_{\|\delta u(\tau)\|^2 \geq 0 \; \forall \tau} d\tau \\ &\geq \int_0^T u^\top(\tau)u(\tau)d\tau = J(u(\cdot)). \end{aligned}$$

Hence, the control (9) achieves the minimum energy.

Minimum energy control (cont.)

The minimum amount of energy can be expressed in terms of the reachability Gramian and the vector η as follows:

$$\begin{aligned}\int_0^T u^\top(\tau)u(\tau)d\tau &= \int_0^T \eta^\top e^{A(T-\tau)}BB^\top e^{A^\top(T-\tau)}\eta d\tau \\ &= \eta^\top\left(\int_0^T e^{A(T-\tau)}BB^\top e^{A^\top(T-\tau)}d\tau\right)\eta = \eta^\top W_r(T)\eta.\end{aligned}$$

If $W_r(T)$ is nonsingular (i.e., if the system is reachable) we can write

$$\eta^\top W_r(T)\eta = \underbrace{\eta^\top W_r(T)}_{\bar{x}^\top} W_r^{-1}(T) \underbrace{W_r(T)\eta}_{\bar{x}} = \bar{x}^\top W_r^{-1}(T)\bar{x}.$$

In other words, if a system is reachable, the minimum energy (in the sense specified above) to reach the state $\bar{x}$ from zero in the interval $[0\ ,\ T]$ is

$$J^* = \bar{x}^\top W_r^{-1}(T)\bar{x},$$

i.e., is a quadratic form of the inverse of the reachability Gramian.

Reachability and controllability for discrete-time linear time-invariant systems

Consider now systems of the form

$$x(k+1) = Ax(k) + Bu(k). \tag{13}$$

There exist strong analogies to the continuous-time case, and some noticeable differences as well. In particular, the definitions hold unchanged.

Definition: reachability and controllability on an interval

Given the states $\bar{x}_1$ and $\bar{x}_2$, and the interval $[k_1, k_2]$, if an input $u : \{k_1, \ldots, k_2 - 1\} \to \mathbb{R}^m$ exists such that

$$x(k_1) = \bar{x}_1 \quad \Longrightarrow \quad x(k_2) = \bar{x}_2$$

then we say that $\bar{x}_2$ is *reachable from* $\bar{x}_1$ on the interval $[k_1, k_2]$ and $\bar{x}_1$ is *controllable to* $\bar{x}_2$ on the same interval.

It is often useful to take $\bar{x}_1 = 0$ and $k_1 = 0$, thus we have the following.

Definition: reachability from zero on an interval

The state vector $\bar{x} \in \mathbb{R}^n$, is said to be *reachable* (from zero) on the interval $[0, K]$ if $\exists u : \{0, \ldots, K - 1\} \to \mathbb{R}^m$ such that

$$x(0) = 0 \quad \Longrightarrow \quad x(K) = \bar{x}$$

Definition: controllability to 0 on an interval

The state vector $\bar{x} \in \mathbb{R}^n$, is said to be *controllable* (to 0) on the interval $[0, K]$ if $\exists u : \{0, \ldots, K - 1\} \to \mathbb{R}^m$ such that

$$x(0) = \bar{x} \quad \Longrightarrow \quad x(K) = 0$$

Definition: reachable set on an interval

Given the time K, the reachable set on $[0, K]$, denoted by $X_r(K)$, consists of all states $\bar{x}$ for which there exists an input $u : \{0, \dots, K-1\} \to \mathbb{R}^m$ that transfers the state from $x(0) = 0$ to $x(K) = \bar{x}$, i.e.,

$$X_r(K) = \{\bar{x} \in \mathbb{R}^n : \bar{x} \text{ is reachable from zero in } [0, K]\}$$

Definition: controllable set on an interval

Given the time K, the controllable (to 0) set on $[0, K]$, denoted by $X_c(K)$, consists of all states $\bar{x}$ for which there exists an input $u : \{0, \dots, K-1\} \to \mathbb{R}^m$ that transfers the state from $x(0) = \bar{x}$ to $x(K) = 0$, i.e.

$$X_c(K) = \{\bar{x} \in \mathbb{R}^n : \bar{x} \text{ is controllable to 0 in } [0, K]\}$$

Definition: reachable set

The reachable set, denoted by X_r, consists of all the states $\bar{x}$ for which there exists a finite K and an input $u : \{0, \dots, K-1\} \to \mathbb{R}^m$ that transfers the state from $x(0) = 0$ to $x(K) = \bar{x}$, i.e.

$$X_r = \{\bar{x} \in \mathbb{R}^n : \bar{x} \text{ is reachable from zero in a finite number of steps}\}$$

The states belonging to X_r are said *reachable*.

Definition: controllable set

The controllable (to 0) set, denoted by X_c, consists of all the states $\bar{x}$ for which there exists a finite K and an input $u : \{0, \dots, K-1\} \to \mathbb{R}^m$ that transfers the state from $x(0) = \bar{x}$ to $x(K) = 0$, i.e.

$$X_c = \{\bar{x} \in \mathbb{R}^n : \bar{x} \text{ is controllable to 0 in a finite number of steps}\}$$

The states belonging to X_c are said *controllable*.

Characterization of $X_r(K)$

Recalling that

$$x(k) = A^k x(0) + \sum_{j=0}^{k-1} A^{k-1-j} B u(j),$$

if the state $\bar{x}$ is reached at time K from $x(0) = 0$, we have

$$x(K) = \bar{x} = \sum_{j=0}^{K-1} A^{K-1-j} B u(j)$$

that can be written as

$$\bar{x} = \underbrace{\begin{bmatrix} B & AB & A^2B & \dots & A^{K-1}B \end{bmatrix}}_{\doteq R_K} \underbrace{\begin{bmatrix} u(K-1) \\ u(K-2) \\ \vdots \\ u(1) \\ u(0) \end{bmatrix}}_{\doteq U_K} = R_K U_K$$

Characterization of $X_r(K)$ (cont.)

Since U_K can be chosen arbitrarily, the set of reachable states in K steps is

$$X_r(K) = \operatorname{im} R_K .$$

Thus:

- as in the continuous-time case, the reachability set $X_r(K)$ is a subspace of the state space;
- contrary to the continuous-time case, the reachability set $X_r(K)$ does depend on the time K.

Since R_{K+1} is obtained from R_K by adding the new columns $A^K B$, the reachability set gets bigger as K increases:

$$X_r(1) \subseteq X_r(2) \subseteq X_r(3) \subseteq \dots$$

This is obvious, because if $\bar{x}$ is reachable in K steps it is reachable in $K+1$ steps as well.
On the other hand, it is easy to show that

$$X_r(n) = X_r(n+1) = \dots$$

where n is the order of the system. Indeed, by the Cayley-Hamilton identity, the columns $A^k B$ with $k \geq n$ are a linear combination of the columns of $B,\ AB,\ A^2B \dots A^{n-1}B$. In other words, the rank of the family of matrices R_K is maximum for $K = n$ and any reachable state can be reached in at most n steps (**either a state can be reached in n steps or it cannot be reached at all**).

An algebraic condition for reachability

By letting $R = R_n$ we have the following

Theorem

The set of all reachable states of (13) is

$$X_r = \operatorname{im} R.$$

As a consequence:

Theorem

The system (13) is reachable if and only if

$$\operatorname{rank} R = n.$$

In general, **for discrete-time systems, controllability and reachability are not equivalent**, as shown by the following examples.

Example

Consider the system

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

We have

$$R = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

Thus, the reachable subspace is the set $\{x = [x_1 \;\; 0]^\top, \forall x_1 \in \mathbb{R}\}$.

However, the set of controllable states is equal to $\mathbb{R}^2$ because, for $u(k) = 0 \;\; \forall k$, and for any $x(0)$, we have

$$x(k) = A^k x(0) = 0, \quad \text{for } k \geq 2.$$

Indeed, A is a *nilpotent* matrix (a matrix M is said to be nilpotent if $M^k = 0$ for some k).

Example

Consider the system

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

We have

$$R = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix},$$

thus the reachable subspace is the set $\{x = [x_1 \ \ 0]^\top, \forall x_1 \in \mathbb{R}\}$.

However, the set of controllable states is equal to $\mathbb{R}^2$ because, for $x(0) = [\alpha \quad \beta]^\top$, by choosing the input $u(0) = -(\alpha + \beta)$ we get

$$x(1) = Ax(0) + Bu(0) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} (-\alpha - \beta) = 0$$

(note that, contrary to the previous example, here we have $A^k = A \neq 0$ for all k).

Controllable set and reachable set (cont.)

To investigate the relationship between X_r and X_c, we need the following

Lemma

If $x \in \operatorname{im} R$, then $Ax \in \operatorname{im} R$, i.e., the reachable subspace X_r is an A-*invariant subspace*.

Proof.
If $x \in \operatorname{im} R$, there exist a vector α such that

$$x = \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix} \alpha$$

which implies

$$Ax = \begin{bmatrix} AB & A^2B & A^3B & \dots & A^nB \end{bmatrix} \alpha.$$

By the Cayley-Hamilton theorem, we have

$$A^n = \beta_0 I + \beta_1 A + \cdots + \beta_{n-1} A^{n-1}$$

for some $\beta_0, \dots, \beta_{n-1}$. Hence, there exists a vector γ such that

$$Ax = \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix} \gamma,$$

thus $Ax \in \operatorname{im} R$. □

Since

$$x(k) = A^k x(0) + \sum_{j=0}^{k-1} A^{k-1-j} Bu(j),$$

it follows that $\bar{x}$ is controllable to zero in K steps if and only if the equation

$$0 = A^K \bar{x} + \sum_{j=0}^{K-1} A^{K-1-j} Bu(j)$$

is satisfied by some input sequence. The previous is equivalent to

$$-A^K \bar{x} = R_K U_K,$$

thus a state $\bar{x}$ belongs to X_c if and only if $A^K \bar{x}$ can be expressed as a linear combination of the columns of R_K:

$$A^K \bar{x} \in \operatorname{im} R_K.$$

Theorem

Given the system $x(k+1) = Ax(k) + Bu(k)$

(i) $X_r \subseteq X_c$ (in words, if a state x is reachable, then it is controllable).

(ii) If $X_r = \mathbb{R}^n$ then $X_c = \mathbb{R}^n$ (in words, if the system is reachable, then it is controllable).

Proof.
(i) If x is reachable, then $x \in \operatorname{im} R$, which is an A-invariant subspace. Hence $A^n x \in \operatorname{im} R$ which implies x is controllable.

(ii) If $\operatorname{im} R = \mathbb{R}^n$, then any $x \in \mathbb{R}^n$ belongs to X_r which is A-invariant. Thus $A^n x \in \operatorname{im} R$ for any x. So any vector x is controllable. □

The previous theorem shows that, for discrete-time systems, reachability implies controllability. If A is nonsingular, however, reachability and controllability are equivalent:

Theorem

Given the system $x(k+1) = Ax(k) + Bu(k)$, **if A is nonsingular**

(i) $X_r = X_c$ (in words, a state x is reachable if and only if it is controllable).

(ii) $X_r = \mathbb{R}^n \iff X_c = \mathbb{R}^n$ (in words, the system is reachable if and only if it is controllable).

Proof.
To prove (i) we only need to show that, if A is nonsingular, any controllable state is reachable as well. Let x be controllable, i.e., $A^n x \in \operatorname{im} R$. If A is nonsingular, then A^{-n} exists, and, by the Cayley-Hamilton theorem, can be written as a linear combination of positive integer powers of A. As a consequence, since $\operatorname{im} R$ is A-invariant:

$$\underbrace{\left(A^{-n}\right)}_{\text{linear combination of positive integer powers of } A} A^n x = x \in \operatorname{im} R,$$

i.e., x is reachable.

The statement (ii) follows obviously. □

Reachability Gramian

Definition

Given the system $x(k+1) = Ax(k) + Bu(k)$, and a time $K > 0$, the matrix

$$W_r(K) = \sum_{i=0}^{K-1} A^{K-(i+1)} BB^\top \left(A^\top\right)^{K-(i+1)}$$

is said to be the *reachability Gramian*.

Clearly, the reachability Gramian is a symmetric positive semi-definite $n \times n$ matrix, for all $K > 0$. Moreover, it is easy to check that

$$W_r(K) = \sum_{i=0}^{K-1} A^i BB^\top \left(A^\top\right)^i = R_K R_K^\top.$$

Reachability Gramian (cont.)

The following result establishes a correspondence between the reachability matrix R and the reachability Gramian $W_r(K)$. The proof is not reported, being similar to that of the continuous-time case.

Theorem

For $K \geq n$, the image of the reachability Gramian $W_r(K)$ is independent of K and is equal to the image of the reachability matrix R:

$$\operatorname{im} W_r(K) = \operatorname{im} R, \quad \forall K \geq n.$$

The following corollary provides a reachability test.

Corollary

The system $x(k+1) = Ax(k) + Bu(k)$ is reachable if and only if

$$\operatorname{rank} W_r(K) = n \quad \forall K \geq n,$$

i.e., the reachability Gramian $W_r(K) = R_K R_K^\top$ is nonsingular for all $K \geq n$.

When a system is reachable, the input sequence that transfers the state from $x(0) = x_0$ to $\bar{x} = x(K)$ can be determined in terms of the reachability Gramian.
Indeed, we need to solve

$$\bar{x} - A^K x_0 = R_K U_K \tag{14}$$

for U_K.
Since the system is reachable, $W_r(K)$ is invertible and, thus, the following sequence is defined

$$U_K = R_K^\top (W_r(K))^{-1} \left(\bar{x} - A^K x_0\right). \tag{15}$$

By substitution, it is immediate to check that it solves (14), because $W_r(K) = R_K R_K^\top$.

Note that $R_K^\top (W_r(K))^{-1} = R_K^\top \left(R_K R_K^\top\right)^{-1}$ is the *Moore-Penrose pseudoinverse* of R_K. Thus, if the solution of (14) is not unique, the solution provided by (15) is the solution of minimum Euclidean norm.

Observability and constructibility for continuous-time linear time-invariant systems

Definitions

The problems of observability and constructibility amount to studying the relationship between the system state and the system output. Consider a linear autonomous continuous-time system with matrices A, B, C, D:

$$\begin{cases} \dot{x}(t) & = & Ax(t) + Bu(t) \\ y(t) & = & Cx(t) + Du(t) \end{cases}$$

Assume that $y(t)$ and $u(t)$ are known in an interval $[0, \tau]$.

Is it possible to uniquely determine the initial state of the system at time $t = 0$?

Is it possible to uniquely determine the final state of the system at time $t = \tau$?

Definitions (cont.)

Definition: observable system on an interval

A system is said to be *observable* in the interval $[0, \tau]$ if, given $u(t)$ and $y(t)$ for $t \in [0, \tau]$, it is possible to *uniquely* determine $x(0)$.

Definition: constructible system on an interval

A system is said to be *constructible* in the interval $[0, \tau]$ if, given $u(t)$ and $y(t)$ for $t \in [0, \tau]$, it is possible to *uniquely* determine $x(\tau)$.

1. Without loss of generality we can assume $D = 0$, because if $D \neq 0$, an auxiliary output $\tilde{y}(t) = y(t) - Du(t)$ can be considered.
2. In the case of continuous-time systems, the observability problem is equivalent to the constructibility problem. Indeed:

$$x(\tau) = e^{A\tau}x(0) + \int_0^t e^{A(\tau-\sigma)}Bu(\sigma)d\sigma$$

If $x(0)$ is uniquely known, then $x(\tau)$ is uniquely known as well.

Conversely, if $x(\tau)$ is uniquely known, since e^{At} is invertible, we have

$$x(0) = e^{-A\tau}x(\tau) - e^{-A\tau}\int_0^t e^{A(\tau-\sigma)}Bu(\sigma)d\sigma = e^{-A\tau}x(\tau) - \int_0^t e^{-A\sigma}Bu(\sigma)d\sigma$$

thus $x(0)$ is uniquely determined.

Unobservable states

The observability problem is conceptually different from the reachability problem discussed earlier. The reachability problem concerns the **existence** of an input that allows us to reach a certain state. The observability deals with the **uniqueness** of the initial state $x(0)$.

In fact, given the same input and observed output, there can be several initial states that are consistent with the observation.

Unobservable states (cont.)

Thus, we study the uniqueness of the solution $x(0)$ to the following equation (recall that we are assuming $D = 0$):

$$y(t) = Ce^{At}x(0) + \underbrace{\int_0^t Ce^{A(t-\sigma)}Bu(\sigma)d\sigma}_{\doteq g(t)} = Ce^{At}x(0) + g(t) \tag{16}$$

where $y(t)$ and $g(t)$ are both known.

Suppose there are two states $\bar{x}_1 \neq \bar{x}_2$ such that, for all $t \in [0, \tau]$

$$\begin{aligned} y(t) &= Ce^{At}\bar{x}_1 + g(t) \\ y(t) &= Ce^{At}\bar{x}_2 + g(t) \end{aligned} \tag{17}$$

that is, two different initial states that produce the same output when the same input is applied.

By subtracting, and letting $\bar{x} = \bar{x}_2 - \bar{x}_1$, we get

$$Ce^{At}\bar{x} = 0 \quad \forall t \in [0, \tau] \tag{18}$$

Unobservable states (cont.)

Then, uniqueness is ensured if and only if there are no vectors $\bar{x} \neq 0$ which satisfy (18). Indeed, if $\bar{x}_1 \neq \bar{x}_2$ satisfy (17), then $\bar{x} = \bar{x}_2 - \bar{x}_1 \neq 0$ satisfies (18). Conversely, if $\bar{x} \neq 0$ satisfies (18) we have

$$\begin{aligned} 0 &= Ce^{At}\bar{x} \\ y(t) &= Ce^{At}x(0) + g(t) \end{aligned}$$

By summing

$$y(t) = Ce^{At}\left(x(0) + \bar{x}\right) + g(t)$$

thus, both the vectors $x(0)$ and $x(0) + \bar{x}$ produce the same output for any $g(\cdot)$ (and thus for any input): then, uniqueness is missing. As a consequence we have the following

Property

A system is observable in the interval $[0, \tau]$ if and only if there are no vectors $\bar{x} \neq 0$ such that $Ce^{At}\bar{x} = 0$ for all $t \in [0, \tau]$.

Unobservable states (cont.)

Definition

A state $\bar{x} \neq 0$ is said to be *unobservable*, or *indistiguishable from zero* in the interval $[0, \tau]$ if $Ce^{At}\bar{x} = 0$ for all $t \in [0, \tau]$.

Thus, a system is observable in $[0, \tau]$ if and only if there exist no unobservable states. It is easy to check that the set of unobservable states in $[0, \tau]$ is a subspace.

Definition

The set $X_{\bar{o}}(\tau)$ of the unobservable states is called the *unobservable subspace*.

Observability matrix

Definition: observability matrix

Given the system

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

the $np \times n$ matrix

$$O = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

is called the *observability matrix*.

Theorem

The unobservable subspace $X_{\bar{o}}(\tau)$ is independent of $\tau > 0$ and is the kernel of the observability matrix:

$$X_{\bar{o}}(\tau) = \ker O = \{x \in \mathbb{R}^n : Ox = 0\}.$$

Proof.
The unobservable subspace $X_{\bar{o}}(\tau)$ is the set of $x \in \mathbb{R}^n$ such that

$$Ce^{At}x = C(I + At + A^2\frac{t^2}{2!} + \ldots)x = \sum_{k=0}^{\infty} C\frac{A^k}{k!}xt^k = 0 \quad \forall t \in [0, \tau] \tag{19}$$

By the principle of identity for power series, this is equivalent to

$$CA^k x = 0 \quad \forall k \geq 0,$$

which is equivalent, by the Cayley-Hamilton theorem, to

$$CA^k x = 0, \quad \forall k = 0, 1, 2, \ldots, n-1.$$

An algebraic condition for observability (cont.)

The last equations can be written in compact form as follows:

$$\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} x = 0 \tag{20}$$

or

$$Ox = 0, \tag{21}$$

which completes the proof. □

An algebraic condition for observability (cont.)

As a consequence we have the following fundamental results.

Theorem

A system is observable if and only if $\ker O = \{0\}$.

Proof.
It follows immediately from the previous theorem, since the observability is equivalent to the non-existence of unobservable states. □

Theorem

A system is observable if and only if

$$\operatorname{rank} O = n,$$

i.e., the observability matrix has full-rank.

Proof.
The matrix O has n columns and np rows, thus its kernel is trivial (i.e., $\ker O = \{0\}$) if and only if the n columns are linearly independent i.e., $\operatorname{rank} O = n$. □

Observation

An observation analogous to that of the reachability is in order. The previous result implies that, if the initial state of a system can be determined, it can be determined in an arbitrarily small observation interval. However, in practice, this is not the case. Each measure $y(t)$, indeed, is certainly affected by noise. To determine the initial state, we need to filter the measured signal and this requires the observation over a period of time that is not infinitesimal.

The unobservable subspace is A-invariant

Lemma

If $x \in \ker O$ then $Ax \in \ker O$, i.e., the unobservable subspace $X_{\bar{o}}$ is A-invariant.

Proof.
Since $x \in \ker O$ we have

$$Ox = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} x = \begin{bmatrix} Cx \\ CAx \\ \vdots \\ CA^{n-1}x \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

As for Ax we have:

$$OAx = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} Ax = \begin{bmatrix} CAx \\ CA^2x \\ \vdots \\ CA^nx \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ CA^nx \end{bmatrix} = 0$$

where the last term is zero by the Cayley-Hamilton theorem. □

Kalman decomposition for observability

If a system is not observable, then $X_{\bar{o}} = \ker O \neq \{0\}$. Let r be the dimension of $X_{\bar{o}}$, take a basis $t_1, \ldots, t_r$, and form the matrix

$$T_{\bar{o}} = [t_1 \ \ldots \ t_r].$$

Then, complete (arbitrarily) a basis of $\mathbb{R}^n$ by choosing $n - r$ linearly independent vectors

$$T_o = [t_{r+1} \ \ldots \ t_n]$$

such that the columns of the matrix

$$T = [T_{\bar{o}} \mid T_o]$$

form a basis of $\mathbb{R}^n$.

Thus any vector $x \in \mathbb{R}^n$ may be written as

$$x = T_{\bar{o}} \hat{x}_{\bar{o}} + T_o \hat{x}_o \ ,$$

where $\hat{x}_{\bar{o}} \in \mathbb{R}^r$ and $\hat{x}_o \in \mathbb{R}^{n-r}$.

By construction, the subspace generated by the columns of $T_{\bar{o}}$ is the unobservable subspace $X_{\bar{o}}$:

$$X_{\bar{o}} = \left\{ x : x = [T_{\bar{o}} \mid T_o] \begin{bmatrix} \alpha \\ 0 \end{bmatrix}, \ \forall \alpha \in \mathbb{R}^r \right\}.$$

We denote the subspace generated by the columns of T_o as the *observable subspace*. This nomenclature is purely conventional, because, if there exist unobservable vectors, the entire observability problem has no solution.

Now we perform a change of basis:

$$x = [T_{\bar{o}} \mid T_o] \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_o \end{bmatrix}$$

or

$$\hat{x} = [T_{\bar{o}} \mid T_o]^{-1} x$$

where

$$\hat{x} = \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_o \end{bmatrix}.$$

The state-space representation of the system becomes

$$\begin{aligned} \begin{bmatrix} \dot{\hat{x}}_{\bar{o}} \\ \dot{\hat{x}}_o \end{bmatrix} &= \begin{bmatrix} A_{\bar{o}} & A_{\bar{o},o} \\ \phi_2 & A_o \end{bmatrix} \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_o \end{bmatrix} + \begin{bmatrix} B_{\bar{o}} \\ B_o \end{bmatrix} u \\ y &= \begin{bmatrix} \phi_1 & C_o \end{bmatrix} \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_o \end{bmatrix}. \end{aligned}$$

Kalman decomposition for observability (cont.)

It can be shown that the blocks ϕ_1 and ϕ_2 are both zero. Assume $u(t) \equiv 0$ and take the initial condition

$$\hat{x}(0) = \begin{bmatrix} \hat{x}_{\bar{o}}(0) \\ 0 \end{bmatrix}.$$

The output is then

$$y(t) = \phi_1 \hat{x}_{\bar{o}}(t) + C_o \hat{x}_o(t).$$

Since the chosen initial condition $\hat{x}(0)$ is unobservabe, i.e., undistinguishable from zero, the output must be identically zero. In particular, for $t = 0$ we have

$$y(0) = \phi_1 \hat{x}_{\bar{o}}(0) + C_o 0 = \phi_1 \hat{x}_{\bar{o}}(0) = 0$$

and being $\hat{x}_{\bar{o}}(0)$ arbitrary, it follows that

$$\phi_1 = 0.$$

On the other hand, if ϕ_2 were not zero, in the same conditions we would have that, at some τ, $\hat{x}_o(\tau) \neq 0$. Such a τ can be seen as new initial time where the new initial condition is not indistinguishable from zero, since it has a non-zero component along the observable subspace. Thus, the output must be different from zero at some subsequent time, which contradicts the unobservability of $\hat{x}(0)$.

As a consequence, the system can be written as

$$\begin{aligned} \begin{bmatrix} \dot{\hat{x}}_{\bar{o}} \\ \dot{\hat{x}}_{o} \end{bmatrix} &= \begin{bmatrix} A_{\bar{o}} & A_{\bar{o},o} \\ 0 & A_o \end{bmatrix} \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_{o} \end{bmatrix} + \begin{bmatrix} B_{\bar{o}} \\ B_o \end{bmatrix} u \\ y &= \begin{bmatrix} 0 & C_o \end{bmatrix} \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_{o} \end{bmatrix} + Du \end{aligned} \tag{22}$$

where the blocks $B_{\bar{o}}$ and B_o do not have any particular property, and D has been introduced back.

In (22) we recognize an observable (proving the observability is left as an exercise) subsystem

$$\Sigma(A_o, B_o, C_o),$$

an unobservable subsystem

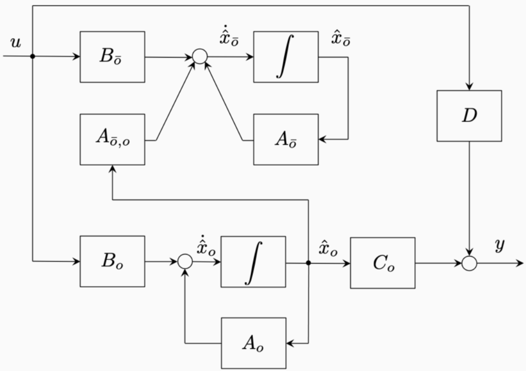
$$\Sigma(A_{\bar{o}}, B_{\bar{o}}, 0),$$

and a block $A_{\bar{o},o}$ representing the action of the observable subsystem on the unobservable one.

Kalman decomposition for observability (cont.)

$$\begin{bmatrix} \dot{\hat{x}}_{\bar{o}} \\ \dot{\hat{x}}_{o} \end{bmatrix} = \begin{bmatrix} A_{\bar{o}} & A_{\bar{o},o} \\ 0 & A_{o} \end{bmatrix} \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_{o} \end{bmatrix} + \begin{bmatrix} B_{\bar{o}} \\ B_{o} \end{bmatrix} u$$

$$y = \begin{bmatrix} 0 & C_{o} \end{bmatrix} \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_{o} \end{bmatrix} + Du$$



$$
\begin{bmatrix} \dot{\hat{x}}_{\bar{o}} \\ \dot{\hat{x}}_{o} \end{bmatrix} = \begin{bmatrix} A_{\bar{o}} & A_{\bar{o},o} \\ 0 & A_{o} \end{bmatrix} \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_{o} \end{bmatrix} + \begin{bmatrix} B_{\bar{o}} \\ B_{o} \end{bmatrix} u
$$

$$
y = \begin{bmatrix} 0 & C_{o} \end{bmatrix} \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_{o} \end{bmatrix} + Du
$$

- the output is not affected by the unobservable subsystem, neither directly nor indirectly through the observable subsystem.

Observations (cont.)

$$\begin{bmatrix} \dot{\hat{x}}_{\bar{o}} \\ \dot{\hat{x}}_o \end{bmatrix} = \begin{bmatrix} A_{\bar{o}} & A_{\bar{o},o} \\ 0 & A_o \end{bmatrix} \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_o \end{bmatrix} + \begin{bmatrix} B_{\bar{o}} \\ B_o \end{bmatrix} u$$

$$y = \begin{bmatrix} 0 & C_o \end{bmatrix} \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_o \end{bmatrix} + Du$$

- the set $\sigma(A)$ of the eigenvalues of A (which, by similarity, are the same of $\hat{A}$) can be partitioned into two subsets:

  $$\sigma(A) = \sigma(A_{\bar{o}}) \cup \sigma(A_o),$$

  where

  - $\sigma(A_{\bar{o}}) = \{\lambda_1, \ldots, \lambda_r\}$ are the *unobservable eigenvalues*, associated with *unobservable modes*, and
  - $\sigma(A_o) = \{\lambda_{r+1}, \ldots, \lambda_n\}$ are the *observable eigenvalues*, associated with *observable modes*.

Observations (cont.)

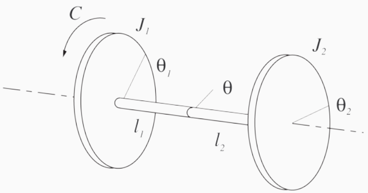
$$\begin{bmatrix} \dot{\hat{x}}_{\bar{o}} \\ \dot{\hat{x}}_o \end{bmatrix} = \begin{bmatrix} A_{\bar{o}} & A_{\bar{o},o} \\ 0 & A_o \end{bmatrix} \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_o \end{bmatrix} + \begin{bmatrix} B_{\bar{o}} \\ B_o \end{bmatrix} u$$

$$y = \begin{bmatrix} 0 & C_o \end{bmatrix} \begin{bmatrix} \hat{x}_{\bar{o}} \\ \hat{x}_o \end{bmatrix} + Du$$

- by recalling that the transfer function is invariant for change of basis, and applying the formula for the inverse of a triangular block matrix, it can be shown that the transfer function depends on the observable subsystem only

$$W(s) = C(sI - A)^{-1}B + D = C_o(sI - A_o)^{-1}B_o + D.$$

Thus, the poles of $W(s)$ are necessarily eigenvalues of A_o.



Example.

Consider two flywheels whose moments of inertia are J_1 and J_2, connected by an elastic joint of stiffness k and length l. A torque $C(t)$ is applied to the first flywheel. Let ϑ_1 and ϑ_2 be the angular positions of the flywheels and ϑ be the angular position of a section of the joint, located at a distance l_1 from the first. The system is governed by the equations

$$\begin{cases} J_1\ddot{\vartheta}_1(t) & = & -k(\vartheta_1(t) - \vartheta_2(t)) + C \\ J_2\ddot{\vartheta}_2(t) & = & -k(\vartheta_2(t) - \vartheta_1(t)) \end{cases} .$$

By assuming a linear angular displacement along the joint, we can write:

$$\vartheta(t) = \frac{l_1}{l}\vartheta_1(t) + \frac{l_2}{l}\vartheta_2(t).$$

Example (cont.)

The state-space representation is

$$\dot{x}(t) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{k}{J_1} & \frac{k}{J_1} & 0 & 0 \\ \frac{k}{J_2} & -\frac{k}{J_2} & 0 & 0 \end{bmatrix} x(t) + \begin{bmatrix} \frac{1}{J_1} \\ 0 \\ 0 \\ 0 \end{bmatrix} u(t)$$

$$y(t) = \begin{bmatrix} \frac{l_1}{l} & \frac{l_2}{l} & 0 & 0 \end{bmatrix} x(t).$$

where $u(t) = C(t)$, $y(t) = \vartheta(t)$ and

$$\begin{aligned} x_1(t) &= \vartheta_1(t) \\ x_2(t) &= \vartheta_2(t) \\ x_3(t) &= \dot{\vartheta}_1(t) \\ x_4(t) &= \dot{\vartheta}_2(t) \end{aligned}$$

Example (cont.)

The observability matrix is:

$$O = \begin{bmatrix} C \\ CA \\ CA^2 \\ CA^3 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{l_1}{l} & \frac{l_2}{l} & 0 & 0 \\ 0 & 0 & \frac{l_1}{l} & \frac{l_2}{l} \\ \frac{k}{l}\left(\frac{l_2}{J_2} - \frac{l_1}{J_1}\right) & \frac{k}{l}\left(\frac{l_1}{J_1} - \frac{l_2}{J_2}\right) & 0 & 0 \\ 0 & 0 & \frac{k}{l}\left(\frac{l_2}{J_2} - \frac{l_1}{J_1}\right) & \frac{k}{l}\left(\frac{l_1}{J_1} - \frac{l_2}{J_2}\right) \end{bmatrix}$$

The matrix has full-rank except when the first column is a multiple of the second (and, as a consequence, the third is a multiple of the fourth).

Example (cont.)

Thus, if

$$\frac{l_1}{J_1} = \frac{l_2}{J_2},$$

the system is not observable. The unobservable subspace has dimension $n - 2 = 2$ and is composed of all the vectors of the form

$$x = \begin{bmatrix} \alpha \\ -\frac{l_1}{l_2}\alpha \\ \beta \\ -\frac{l_1}{l_2}\beta \end{bmatrix}, \quad \alpha, \beta \in \mathbb{R}.$$

Those vectors are indistinguishable from zero.

Observability Gramian

Definition

Given the system

$$\begin{cases} \dot{x}(t) & = & Ax(t) + Bu(t) \\ y(t) & = & Cx(t) + Du(t) \end{cases}$$

and a time $T > 0$, the matrix

$$W_o(T) = \int_0^T e^{A^\top \tau} C^\top C e^{A\tau} d\tau$$

is said to be the *observability Gramian*.

Clearly, the observability Gramian is a symmetric positive semi-definite $n \times n$ matrix, for all $T > 0$.

Observability Gramian (cont.)

The following result establishes a correspondence between the observability matrix O and the observability Gramian $W_o(T)$.

Theorem

The null space of the observability Gramian $W_o(T)$ is independent of $T > 0$ and is equal to the null space of the observability matrix O:

$$\ker W_o(T) = \ker O, \quad \forall T > 0.$$

Proof.

We first show that if $x \in \ker O$ then $x \in \ker W_o(T)$:

$$x \in \ker O \quad \Longrightarrow \quad Ox = 0 \quad \Longrightarrow \quad CA^k x = 0 \quad k = 0, \ldots, n-1.$$

The last equality holds also for $k > n-1$ by the Cayley-Hamilton theorem. Thus:

$$Ce^{At}x = C \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k x = 0, \quad \forall t > 0.$$

As a consequence:

$$W_o(T)x = \int_0^T e^{A^\top \tau} C^\top C e^{A\tau} x d\tau = 0, \quad \forall T > 0$$

which means that $x \in \ker W_o(T)$ for all $T > 0$.

Observability Gramian (cont.)

Conversely, let $x \in \ker W_o(T)$. Then:

$$x^\top \underbrace{W_o(T)x}_{=0} = \int_0^T \|Ce^{A\tau}x\|^2 d\tau = 0$$

which implies that

$$Ce^{At}x = 0, \quad \forall t \in [0, \;\; T].$$

By taking the derivatives w.r.t. t and evaluating at $t = 0$ we get

$$Cx = CAx = \cdots = CA^k x = 0, \quad \forall k > 0 \quad \Longrightarrow \quad Ox = 0.$$

Thus, $x \in \ker O$. □

Observability Gramian (cont.)

The following corollary provides an observability test.

Corollary

The system

$$\begin{cases} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{cases}$$

is observable if and only if

$$\operatorname{rank} W_o(T) = n \quad \forall T > 0,$$

i.e., the observability Gramian is nonsingular for all $T > 0$.

Observability Gramian (cont.)

If the system is observable, the problem of recovering the initial state $x(0) = x_0$ from the knowledge of the output $y(t)$ and the input $u(t)$ in the interval $[0, \ T]$ can be solved by means of the Gramian, as shown next.

Lemma

If the system is observable,

$$x_0 = W_o^{-1}(T) \int_0^T e^{A^\top t} C^\top \tilde{y}(t) dt \tag{23}$$

where

$$\tilde{y}(t) \doteq y(t) - \left[\int_0^t C e^{A(t-\tau)} B u(\tau) d\tau + D u(t) \right]. \tag{24}$$

Proof.
The output $y(t)$ of the system at time t, starting from $x(0) = x_0$, and applying the input $u(\cdot)$ is

$$y(t) = Ce^{At}x_0 + \left[\int_0^t Ce^{A(t-\tau)}Bu(\tau)d\tau + Du(t)\right],$$

thus, from (24), we have

$$Ce^{At}x_0 = \tilde{y}(t).$$

Premultiplying by $e^{A^\top t}C^\top$, we get:

$$e^{A^\top t}C^\top Ce^{At}x_0 = e^{A^\top t}C^\top \tilde{y}(t).$$

Finally, integrating over $[0, \quad T]$ we get

$$\underbrace{\int_0^T e^{A^\top t}C^\top Ce^{At}dt}_{W_o(T)}\, x_0 = W_o(T)x_0 = \int_0^T e^{A^\top t}C^\top \tilde{y}(t)dt.$$

When the system is observable, (23) is the unique solution for x_0. □

Observability and constructibility for discrete-time linear time-invariant systems

Observability

The observability problem for discrete-time systems is stated exactly as in the continuous-time case.

Given the system

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) \end{aligned} \tag{25}$$

the output at step $k > 0$ is

$$y(k) = CA^k x(0) + \underbrace{\sum_{h=0}^{k-1} CA^{k-h-1} Bu(h)}_{\doteq g(k)} = CA^k x(0) + g(k).$$

Now, assuming to perform T observations and recalling that $y(0) = Cx(0)$, we get T equations:

$$\begin{aligned} y(0) &= Cx(0) \\ y(1) &= CAx(0) + g(1) \\ &\vdots \\ y(T-1) &= CA^{T-1}x(0) + g(T-1) \end{aligned}$$

Observability (cont.)

or, in compact form:

$$Y_T - G_T = O_T x(0), \tag{26}$$

where

$$Y_T = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(T-1) \end{bmatrix}, \; G_T = \begin{bmatrix} 0 \\ g(1) \\ \vdots \\ g(T-1) \end{bmatrix}, \; O_T = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{T-1} \end{bmatrix}.$$

The system (26) has a unique solution for $x(0) \in \mathbb{R}^n$ if and only if O_T has trivial kernel, or, equivalently, it has n linearly independent columns.

Notice that, by the Cayley-Hamilton theorem:

$$\operatorname{rank} O_n = \operatorname{rank} O_{n+1} = \dots$$

thus, either $x(0)$ can be uniquely determined from $T = n$ observations or it cannot be uniquely determined at all. Therefore, the observability depends on the rank of the observability matrix $O \doteq O_n$.

Theorem

The system (25) is observable if and only if

$$\operatorname{rank} O = n,$$

i.e., the observability matrix has full-rank.

As for the continuous-time case, it can be shown that the set $X_{\bar{o}}$ of the unobservable states is a subspace, and precisely it is the kernel of the observability matrix O. Moreover, the Kalman decomposition can be obtained by referring the system to an arbitrarily completed basis of $X_{\bar{o}}$.

Constructibility

Since

$$x(k) = A^k x(0) + \sum_{h=0}^{k-1} A^{k-h-1} B u(h), \tag{27}$$

it is easy to show that observability implies constructibility (as in the continuous-time case). Indeed, if $x(0)$ is uniquely determined from $y(\cdot)$ and $u(\cdot)$, the previous equation provides $x(k)$.

Contrary to the continuous-time case, the reverse is not true: in other words, a discrete-time system can be constructible but not observable. As an example, consider the system

$$\begin{aligned} \begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} \\ y(k) &= \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}. \end{aligned}$$

As the output is identically zero, the system is clearly not observable. It is however constructible because the state after two steps is uniquely determined: $x(k) = 0, \ \forall k \geq 2$.

In the discrete-time case, observability and constructibility are equivalent if the system is *reversible*, i.e., the matrix A is invertible. In that case, the matrix A^{-k} exists for all $k > 0$, therefore the equation

$$x(k) = A^k x(0) + \sum_{h=0}^{k-1} A^{k-h-1} B u(h),$$

can be solved for $x(0)$:

$$x(0) = A^{-k} \left[x(k) - \sum_{h=0}^{k-1} A^{k-h-1} B u(h) \right].$$

Definition

Given the system

$$\begin{cases} x(k+1) & = & Ax(k) + Bu(k) \\ y(k) & = & Cx(k) + Du(k) \end{cases}$$

and a time $K > 0$, the matrix

$$W_o(K) = \sum_{i=0}^{K-1} \left(A^\top\right)^i C^\top C A^i$$

is said to be the *observability Gramian*.

Clearly, the observability Gramian is a symmetric positive semi-definite $n \times n$ matrix, for all $K > 0$. Moreover, it is easy to check that

$$W_o(K) = O_K^\top O_K.$$

Observability Gramian (cont.)

The following result establishes a correspondence between the observability matrix O and the observability Gramian $W_o(K)$. The proof is not reported, being similar to that of the continuous-time case.

Theorem

For $K \geq n$, the null space of the observability Gramian $W_o(K)$ is independent of K and is equal to the null space of the observability matrix O:

$$\ker W_o(K) = \ker O, \quad \forall K \geq n.$$

The following corollary provides an observability test.

Observability Gramian (cont.)

Corollary

The system

$$\begin{cases} x(k+1) &= & Ax(k) + Bu(k) \\ y(k) &= & Cx(k) + Du(k) \end{cases}$$

is observable if and only if

$$\operatorname{rank} W_o(K) = n$$

for some (and consequently for all) $K \geq n$.

Observability Gramian (cont.)

When a system is observable, and $K \geq n$, the state $x(0) = x_0$ is given by

$$x_0 = W_o^{-1}(K) O_K^\top \left[Y_K - M_K U_K \right],$$

where:

$$Y_K = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(K-1) \end{bmatrix}, \quad U_K = \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(K-1) \end{bmatrix}$$

and

$$M_K = \begin{bmatrix} D & 0 & \dots & 0 \\ CA^0B & D & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ CA^{K-2}B & \dots & CA^0B & D \end{bmatrix}.$$

Observability Gramian (cont.)

Indeed, the output at time k is given by

$$y(k) = CA^k x_0 + \sum_{h=0}^{k-1} CA^{k-h-1} Bu(h) + Du(k),$$

which can be evaluated for $k = 0, \dots, K-1$, leading to a set of equations that can easily be arranged as

$$Y_K = O_K x_0 + M_K U_K$$

and, equivalently

$$O_K x_0 = Y_K - M_K U_K.$$

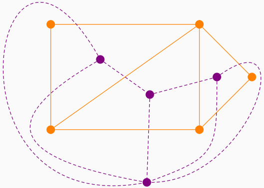
Since the system is observable and $K \geq n$, O_K has full-column rank (n). Thus, the previous linear system of equations admits the unique solution:

$$x_0 = \underbrace{\left(O_K^\top O_K\right)^{-1}}_{W_o^{-1}(K)} O_K^\top \left(Y_K - M_K U_K\right),$$

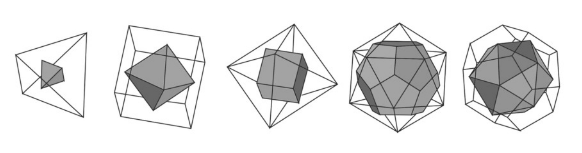
obtained via the pseudo-inverse.

Duality

Dual systems



(a) Dual graphs.



(b) Duality of Platonic solids.

Duality is a fundamental aspect of mathematics and arises in numerous different areas (see Luenberger (1992) for some examples).

Duality between linear systems is a useful tool. In the following, we introduce some essential facts about the duality of linear systems.

Dual systems (cont.)

Given the *primal* system

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \qquad \text{(primal)}$$

where $u(t) \in \mathbb{R}^m$, $x(t) \in \mathbb{R}^n$ and $y(t) \in \mathbb{R}^p$, let the *dual* system be defined as

$$\begin{aligned} \dot{z}(t) &= A^\top z(t) + C^\top v(t) \\ w(t) &= B^\top z(t) + D^\top v(t) \end{aligned} \qquad \text{(dual)}$$

An easy way to get the dual system matrices is taking the transpose as follows:

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \xrightarrow{\quad\top\quad} \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]^\top = \left[\begin{array}{c|c} A^\top & C^\top \\ \hline B^\top & D^\top \end{array}\right]$$

As a consequence of the previous definition:

- The dual system has the same number n of state variables of the primal: $z(t) \in \mathbb{R}^n$;
- the dual system has m outputs (i.e., the number of inputs of the primal): $w(t) \in \mathbb{R}^m$;
- the dual system has p inputs (i.e., the number of outputs of the primal): $v(t) \in \mathbb{R}^p$.

Properties

The following properties are obvious or easy to prove:

- The dual of $\Sigma(A, B, C, D)$ is $\Sigma(A_*, B_*, C_*, D_*) = \Sigma(A^\top, C^\top, B^\top, D^\top)$.
- The dual of the dual is the primal.
- The reachability matrix of the dual is:
$$R_* = O^\top$$
i.e., the transpose observability matrix of the primal.
- The observability matrix of the dual is:
$$O_* = R^\top$$
i.e., the transpose reachability matrix of the primal.
- The transfer function of the dual is
$$W_*(s) = C_*(sI - A_*)^{-1}B_* + D_* = B^\top(sI - A^\top)^{-1}C^\top + D^\top = [W(s)]^\top$$
i.e., the transpose transfer function of the primal.
- In the single-input single output (SISO) case, the primal and the dual have the same transfer function.

Duality reachability/observability

Theorem

Given the systems

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \qquad \text{(primal)}$$

and

$$\begin{aligned} \dot{z}(t) &= A^\top z(t) + C^\top v(t) \\ w(t) &= B^\top z(t) + D^\top v(t) \end{aligned} \qquad \text{(dual)}$$

- The primal is reachable $\iff$ the dual is observable.
- The primal is observable $\iff$ the dual is reachable.

Proof.
It is immediate recalling that $R_* = O^\top$ and $O_* = R^\top$. □

Theorem

The system

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

is observable if and only if

$$\operatorname{rank}\begin{bmatrix}\lambda I - A\\ C\end{bmatrix} = n, \quad \forall \lambda \in \mathbb{C}.$$

Proof.
It follows, by duality, from the PBH test for reachability. □

Observe that:

- it is sufficient to check $\lambda \in \sigma(A)$, i.e., the set of eigenvalues of A;
- the matrix $\begin{bmatrix}\lambda I - A\\ C\end{bmatrix}$ has a rank drop for each $\lambda \in \sigma(A_{\bar{o}})$ i.e., for each unobservable eigenvalue.

Kalman canonical decomposition

Kalman canonical decomposition

Any system $\Sigma(A, B, C, D)$ can be put in a form, called *Kalman canonical form*, that takes into account both the reachability and observability properties.

The reachable and unobservable subspaces are, respectively:

$$X_r = \operatorname{im} R \qquad \text{and} \qquad X_{\bar{o}} = \ker O.$$

Let define the following subspaces (where $\oplus$ denotes the direct sum, i.e., $V_1 \oplus V_2 = V_3$ means that $V_1 + V_2 = V_3$, and $V_1 \cap V_2 = \{0\}$):

- X_1 is the intersection of the reachable and unobservable subspaces, i.e.

$$X_1 \doteq X_r \bigcap X_{\bar{o}}$$

- X_2 is a completion of X_1 in X_r i.e.

$$X_1 \oplus X_2 = X_r$$

- X_3 is a completion of X_1 in $X_{\bar{o}}$ i.e.

$$X_1 \oplus X_3 = X_{\bar{o}}$$

- X_4 is a completion of the previous subspaces in $\mathbb{R}^n$ i.e.

$$X_1 \oplus X_2 \oplus X_3 \oplus X_4 = \mathbb{R}^n$$

Kalman canonical decomposition (cont.)

Then, by defining $X_{\bar{r}} \doteq X_3 \oplus X_4$ and $X_o \doteq X_2 \oplus X_4$, we have:

$$\begin{aligned} X_r &= X_1 \oplus X_2 \\ X_{\bar{r}} &= X_3 \oplus X_4 \\ X_{\bar{o}} &= X_1 \oplus X_3 \\ X_o &= X_2 \oplus X_4. \end{aligned}$$

Let T_1, T_2, T_3, T_4 be bases of X_1, X_2, X_3, X_4 respectively. Clearly, the matrix

$$T = [T_1 \mid T_2 \mid T_3 \mid T_4]$$

is invertible. Any $x \in \mathbb{R}^n$ can be written as

$$x(t) = T_1\hat{x}_1(t) + T_2\hat{x}_2(t) + T_3\hat{x}_3(t) + T_4\hat{x}_4(t)$$

where the dimension of the vector $\hat{x}_i(t)$ is the same as that of X_i.

By performing a change of basis according to $x(t) = T\hat{x}(t)$, we get

$$\frac{d}{dt}\begin{bmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ \hat{x}_3(t) \\ \hat{x}_4(t) \end{bmatrix} = \begin{bmatrix} A_1 & A_{12} & A_{13} & A_{14} \\ 0 & A_2 & 0 & A_{24} \\ 0 & 0 & A_3 & A_{34} \\ 0 & 0 & 0 & A_4 \end{bmatrix}\begin{bmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ \hat{x}_3(t) \\ \hat{x}_4(t) \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{bmatrix} u(t)$$

$$y(t) = \begin{bmatrix} 0 & C_2 & 0 & C_4 \end{bmatrix}\begin{bmatrix} \hat{x}_1(t) \\ \hat{x}_2(t) \\ \hat{x}_3(t) \\ \hat{x}_4(t) \end{bmatrix} + Du(t)$$

Notice that all the zeros in the previous representation are structural. In the following we prove that the block A_{21} is null; similar proofs hold for the other blocks.

Kalman canonical decomposition (cont.)

Take the matrix

$$T^{-1} = \left[\begin{array}{c} H_1 \\ \hline H_2 \\ \hline H_3 \\ \hline H_4 \end{array}\right]$$

partitioned in accordance to T. We have

$$T^{-1}T = \left[\begin{array}{c} H_1 \\ \hline H_2 \\ \hline H_3 \\ \hline H_4 \end{array}\right] \left[\begin{array}{c|c|c|c} T_1 & T_2 & T_3 & T_4 \end{array}\right] = \begin{bmatrix} H_1 T_1 & H_1 T_2 & H_1 T_3 & H_1 T_4 \\ H_2 T_1 & H_2 T_2 & H_2 T_3 & H_2 T_4 \\ H_3 T_1 & H_3 T_2 & H_3 T_3 & H_3 T_4 \\ H_4 T_1 & H_4 T_2 & H_4 T_3 & H_4 T_4 \end{bmatrix}$$

Since $T^{-1}T = I$, we get in particular that $H_2 T_1 = 0$.

On the other hand, we have

$$T^{-1}AT = \begin{bmatrix} H_1 \\ \hline H_2 \\ \hline H_3 \\ \hline H_4 \end{bmatrix} A \left[\begin{array}{c|c|c|c} T_1 & T_2 & T_3 & T_4 \end{array} \right]$$

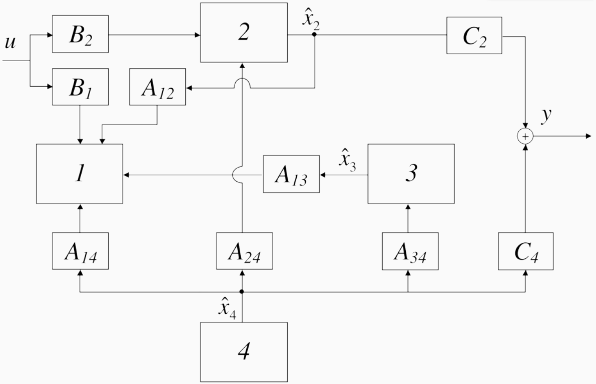
thus $A_{21} = H_2AT_1$.

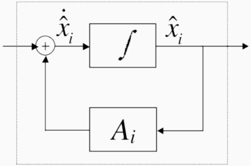
Now observe that each column of T_1 belongs to $X_1 = X_r \bigcap X_{\bar{o}}$ and since X_r and $X_{\bar{o}}$ are A-invariant, each column of AT_1 belongs to X_1 too. Thus AT_1 may be written as $AT_1 = T_1M$ for some matrix M. As a consequence:

$$A_{21} = H_2AT_1 = \underbrace{H_2T_1}_{=0} M = 0.$$

Kalman canonical decomposition (cont.)

The obtained representation corresponds (for $D = 0$) to the following scheme (left), where each numbered block is a subsystem of the type shown on the right.





Kalman canonical decomposition (cont.)

We can recognize the following subsystems:

- $\Sigma_1(A_1)$ is the reachable and unobservable subsystem;
- $\Sigma_2(A_2)$ is the reachable and observable subsystem;
- $\Sigma_3(A_3)$ is the unreachable and unobservable subsystem;
- $\Sigma_4(A_4)$ is the unreachable and observable subsystem.

Theorem

The transfer function depends on the reachable and observable subsystem only, i.e.

$$W(s) = C(sI - A)^{-1}B + D = C_2(sI - A_2)^{-1}B_2 + D$$

The proof of the theorem is rather easy, by explicit computation. This result is not surprising, since unreachable components are not affected by the input, while unobservable components do not give any contribution to the output.

Analogously, the impulse response matrix is

$$W(t) = Ce^{At}B + D\delta(t) = C_2e^{A_2t}B_2 + D\delta(t).$$

Canonical decomposition and BIBO stability

The Kalman decomposition provides further insight on the BIBO stability as related to the asymptotic stability, as shown by the following theorem.

Theorem

A system is BIBO stable if and only if the reachable and observable subsystem is asymptotically stable. If the whole system is reachable and observable, then asymptotic stability is equivalent to BIBO stability.

We will not formally prove the theorem, but the intuition is as follows. On the one hand, BIBO stability depends on the transfer function, which is not affected by non-reachable and/or non-observable parts (if any). Conversely, reachable unstable (or marginally stable) modes can become arbitrarily large due to a bounded input and, if they are observable, they will produce an arbitrarily large output.

References

References

Luenberger, D. G. (1992). A double look at duality. *IEEE Transactions on Automatic Control*, 37(10):1474–1482.

322MI –Spring 2023

Lecture 4

Structural properties and special forms

END