OBSERVABILITY AND CONTROLLABILITY

5.1: Continuous-time observability: Where am I?

- We describe dual ideas called observability and controllability.
- Both have precise (binary) mathematical descriptions, but we need to be careful in interpreting the result.
- We develop some other techniques to help quantify the concepts.

Continuous-time observability

- If a system is observable, we can determine the initial condition of the state vector \( x(0) \) via processing the input to the system \( u(t) \) and the output of the system \( y(t) \).
- Since we can simulate the system if we know \( x(0) \) and \( u(t) \) this also implies that we can determine \( x(t) \) for \( t \geq 0 \). e.g., for LTI,
  \[
  x(t) = e^{At}x(0) + \int_0^t e^{A(t-\tau)}Bu(\tau)\,d\tau.
  \]
- Consider the LTI SISO system LCCODE
  \[
  \ddot{y}(t) + a_1 \dot{y}(t) + a_2 \dot{y}(t) + a_3 y(t) = b_0 \ddot{u}(t) + b_1 \dot{u}(t) + b_2 \dot{u}(t) + b_3 u(t).
  \]
- If we have a realization of this system in state-space form
  \[
  \dot{x}(t) = Ax(t) + Bu(t) \\
  y(t) = Cx(t) + Du(t),
  \]
and we have initial conditions \( y(0), \dot{y}(0), \ddot{y}(0) \), how do we find \( x(0) \)?

\[
y(0) = Cx(0) + Du(0) \\
\dot{y}(0) = C( Ax(0) + Bu(0) ) + D\dot{u}(0) \\
= CAx(0) + CBu(0) + D\dot{u}(0) \\
\ddot{y}(0) = CA^2x(0) + CABu(0) + CB\dot{u}(0) + D\ddot{u}(0).
\]

- In general,

\[
y^{(k)}(0) = CA^kx(0) + CA^{k-1}Bu(0) + \cdots + CBu^{(k-1)}(0) + Du^{(k)}(0),
\]

or,

\[
\begin{bmatrix}
  y(0) \\
  \dot{y}(0) \\
  \ddot{y}(0)
\end{bmatrix}
= \begin{bmatrix}
  C \\
  CA \\
  CA^2
\end{bmatrix}
\begin{bmatrix}
  x(0) + \sum_{i=0}^{k-1}
  \begin{bmatrix}
    D & 0 & 0 \\
    CB & D & 0 \\
    CAB & CB & D
  \end{bmatrix}
  \begin{bmatrix}
    u^{(i)}(0) \\
    \dot{u}^{(i)}(0) \\
    \ddot{u}^{(i)}(0)
  \end{bmatrix}
\end{bmatrix},
\]

where \( \mathcal{T} \) is a (block) “Toeplitz matrix”.

- Thus, if \( \mathcal{O}(C, A) \) is invertible, then

\[
x(0) = \mathcal{O}^{-1} \left\{ \begin{bmatrix}
  y(0) \\
  \dot{y}(0) \\
  \ddot{y}(0)
\end{bmatrix} - \mathcal{T} \begin{bmatrix}
  u(0) \\
  \dot{u}(0) \\
  \ddot{u}(0)
\end{bmatrix} \right\}.
\]

- We say that \( \{C, A\} \) is an observable pair if \( \mathcal{O} \) is nonsingular.

**CONCLUSION:** For a SISO system, if \( \mathcal{O} \) is nonsingular, then we can determine/estimate the initial state of the system \( x(0) \) using only \( u(t) \) and \( y(t) \) (and therefore, we can estimate \( x(t) \) for all \( t \geq 0 \)).

**EXTENSION:** For a MIMO system, if \( \mathcal{O} \) is full rank, then we can determine/estimate the initial state of the system \( x(0) \) using only \( u(t) \) and \( y(t) \) (and therefore, we can estimate \( x(t) \) for all \( t \geq 0 \)).
EXAMPLE: Observability canonical form:

\[
\dot{x}(t) = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
-a_3 & -a_2 & -a_1
\end{bmatrix} x(t) + \begin{bmatrix}
\beta_1 \\
\beta_2 \\
\beta_3
\end{bmatrix} u(t)
\]

\[
y(t) = \begin{bmatrix}
1 & 0 & 0
\end{bmatrix} x(t).
\]

Then

\[
\mathcal{O} = \begin{bmatrix}
C \\
CA \\
CA^2
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 \\
0 & \cdots & 0 \\
0 & 0 & 1
\end{bmatrix} = I_n.
\]

This is why it is called observability form!

EXAMPLE: Two unobservable networks

In the first, if \( u(t) = 0 \) then \( y(t) = 0 \quad \forall \ t \). Cannot determine \( x(0) \).

In the second, if \( u(t) = 0, x_1(0) \neq 0 \) and \( x_2(0) = 0 \), then \( y(t) = 0 \) and we cannot determine \( x_1(0) \). [circuit redrawn for \( u(t) = 0 \)].

Observers

An observer is a device that has as inputs \( u(t) \) and \( y(t) \)—the input and output of a linear system. The output of the observer is the (estimated) state of the linear system.
- The observer "observes" the internal state $x$ (estimated as $\hat{x}$) from external signals $u$ and $y$.

- Note that our equations yield an observer:

- Later, we’ll design more practical observers that don’t use differentiators.
5.2: Continuous-time controllability: Can I get there from here?

- Can we generate an input \( u(t) \) to set an initial condition quickly?

\[
\dot{x}(t) = Ax(t) + Bu(t) \\
y(t) = Cx(t) + Du(t).
\]

- If \( u(t) = \delta(t) \) and \( x(0^-) = 0 \), then

\[
X(s) = (sI - A)^{-1}BU(s) = (sI - A)^{-1}B.
\]

- So, via the Laplace initial-value theorem,

\[
x(0^+) = \lim_{s \to \infty} sX(s) \\
= \lim_{s \to \infty} s(sI - A)^{-1}B \\
= \lim_{s \to \infty} \left(I - \frac{A}{s}\right)^{-1}B \\
= B.
\]

- Thus, an impulse input brings the state to \( B \) from \( 0 \).

- What if \( u(t) = \delta^{(k)}(t) \)?

- Then

\[
X(s) = (sI - A)^{-1}Bs^k = \frac{1}{s} \left(I - \frac{A}{s}\right)^{-1}Bs^k \\
= \frac{1}{s} \left(I + \frac{A}{s} + \frac{A^2}{s^2} + \ldots\right) Bs^k
\]

holds for large \( s \)

\[
= Bs^{k-1} + ABs^{k-2} + A^2Bs^{k-3} + \cdots + \frac{A^kB}{s} + \frac{A^{k+1}B}{s^2} + \cdots
\]

- The first terms are impulsive: they have zero value for \( t > 0 \).
Thus,

\[
x(0^+) = \lim_{s \to \infty} s \left( \frac{A^k B}{s} + \frac{A^{k+1} B}{s^2} + \cdots \right)
\]

\[= A^k B.
\]

So, if \( u(t) = \delta^{(k)}(t) \) then \( x(0^+) = A^k B \).

Now, consider the input

\[u(t) = g_1 \delta(t) + g_2 \dot{\delta}(t) + \cdots g_n \delta^{(n)}(t).
\]

Since \( x(0^-) = 0 \), \( x(0^+) = g_1 B + g_2 A B + \cdots + g_n A^{n-1} B \), or

\[
x(0^+) = \begin{bmatrix}
B & AB & \cdots & A^{n-1} B
\end{bmatrix}
\begin{bmatrix}
g_1 \\
g_2 \\
g_3 \\
\end{bmatrix}

\]

where \( C \) is called the “controllability matrix.”

**CONCLUSION:** For a SISO system, if \( C \) is nonsingular, then there is an impulsive input \( u \) such that \( x(0^+) \) is any desired vector if \( x(0^-) = 0 \).

**EXTENSION:** For a MIMO system, if \( C \) is full rank, then there is an impulsive input \( u \) such that \( x(0^+) \) is any desired vector if \( x(0^-) = 0 \).

In fact, we may use

\[
u(t) = \sum_{i=0}^{n-1} g_i \delta^{(i)}(t)
\]

where

\[
\begin{bmatrix}
g_1 \\
g_2 \\
g_3 \\
\end{bmatrix}
 = \begin{bmatrix}
B & AB & \cdots & A^{n-1} B
\end{bmatrix}^{-1} x_d
\]

where \( x_d \) is the desired \( x(0^+) \) vector.
If $C$ is nonsingular, we say $\{A, B\}$ is a controllable pair and the system is controllable.

**EXAMPLE:** Controllability canonical form:

\[
\dot{x}(t) = \begin{bmatrix} 0 & 0 & -a_3 \\ 1 & 0 & -a_2 \\ 0 & 1 & -a_1 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u(t)
\]

\[
y(t) = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 \end{bmatrix} x(t).
\]

Then

\[
C = \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cdots & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_n.
\]

This is why it is called controllability form!

If a system is controllable, we can instantaneously move the state from any known state to any other state, using impulse-like inputs.

Later, we’ll see that smooth inputs can effect the state transfer (not instantaneously, though!).

**DUALITY:** $\{A, B, C, D\}$ controllable $\iff \{A^T, C^T, B^T, D^T\}$ observable.

**EXAMPLE:** Two uncontrollable networks.
In the first one, if $x(0) = 0$ then $x(t) = 0 \quad \forall \ t$. Cannot influence state!

In the second one, if $x_1(0) = x_2(0)$ then $x_1(t) = x_2(t) \quad \forall \ t$. Cannot independently alter state.

**Diagonal systems, controllability and observability**

Recall the diagonal form

$$
\begin{align*}
\dot{x}(t) &= \begin{bmatrix}
\lambda_1 & 0 \\
0 & \lambda_2 \\
& \cdots \\
0 & \lambda_n
\end{bmatrix} x(t) + \begin{bmatrix}
\gamma_1 \\
\gamma_2 \\
& \cdots \\
\gamma_n
\end{bmatrix} u(t) \\
y(t) &= \begin{bmatrix}
\delta_1 & \delta_2 & \cdots & \delta_n
\end{bmatrix} x(t) + \begin{bmatrix}
0
\end{bmatrix} u(t).
\end{align*}
$$

![Diagram](image)

When controllable? When observable?

$$
O = \begin{bmatrix}
C \\
CA \\
& \vdots \\
& CA^{n-1}
\end{bmatrix} = \begin{bmatrix}
\delta_1 & \delta_2 & \cdots & \delta_n \\
\lambda_1 \delta_1 & \lambda_2 \delta_2 & \cdots & \lambda_n \delta_n \\
& \cdots \\
& \lambda_1^{n-1} \delta_1 & \lambda_2^{n-1} \delta_2 & \cdots & \lambda_n^{n-1} \delta_n
\end{bmatrix}
$$
\[
\begin{bmatrix}
1 & 1 & \cdots & 1 \\
\lambda_1 & \lambda_2 & \cdots & \lambda_n \\
\vdots \\
\lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1}
\end{bmatrix}
\]

\[V\]

- Singular?

\[
\det\{\mathcal{O}\} = (\delta_1 \cdots \delta_n) \det\{\mathcal{V}\} = (\delta_1 \cdots \delta_n) \prod_{i<j}(\lambda_j - \lambda_i).
\]

CONCLUSION: Observable \iff \lambda_i \neq \lambda_j, i \neq j and \delta_i \neq 0 \ for \ i = 1, \cdots, n.

- If \(\lambda_1 = \lambda_2\) then not observable. Can only “observe” the sum \(x_1 + x_2\).
- If \(\delta_k = 0\) then cannot observe mode \(k\).
- What about controllability? Use duality and switch \(\delta\)s and \(\gamma\)s.

CONCLUSION: Controllable \iff \lambda_i \neq \lambda_j, i \neq j and \gamma_i \neq 0 \ for \ i = 1, \cdots, n.

- If \(\lambda_1 = \lambda_2\) then not controllable. Can only “control” the sum \(x_1 + x_2\).
- If \(\gamma_k = 0\) then cannot control mode \(k\).
5.3: Discrete-time controllability and observability

Discrete-time controllability

- Similar concept for discrete-time.
- Consider the problem of driving a system to some arbitrary state \( x[n] \)

\[
x[k + 1] = Ax[k] + Bu[k]
\]

\[
x[1] = Ax[0] + Bu[0]
\]

\[
\]

\[
\]

\[
\vdots
\]

\[
x[n] = A^n x[0] + \underbrace{B A B A^2 B \cdots A^{n-1} B}_C \begin{bmatrix} u[n-1] \\ \vdots \\ u[0] \end{bmatrix}.
\]

- Which leads to

\[
\begin{bmatrix} u[n-1] \\ \vdots \\ u[0] \end{bmatrix} = C^{-1} [x[n] - A^n x[0]].
\]

If \( C \) has no inverse (\( \det(C) = 0 \), \( C \) is not full-rank) then these control signals don’t exist. In that case, the input is only partially effective in influencing the state.

- If \( C \) is full-rank, then the input can move the system to any arbitrary state for \( x[0] \).

**NOTE I:** State transition is not instantaneous. Takes \( n \) time steps.
NOTE II: In continuous-time, we used input \( u(t) = g_0 \delta(t) + g_1 \dot{\delta}(t) + \cdots \), a signal we could only approximate in practice. Here, the input is a perfectly good input signal.

**Discrete-time reachability**

- In the literature, there are three different controllability definitions:
  1. Transfer any state to any other state.
  2. Transfer any state to zero, called *controllability to the origin*.
  3. Transfer the zero state to any state, called *controllability from the origin*, or *reachability*.

- In continuous time, because \( e^{At} \) is nonsingular, the three definitions are equivalent.
- In discrete time, if \( A \) is nonsingular, the three definitions are also equivalent.
- However, if \( A \) is singular, (1) and (3) are equivalent but not (2) and (3).

**EXAMPLE:**

\[
x[k + 1] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} x[k] + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} u[k].
\]

- Its controllability matrix has rank 0 and the equation is not controllable in (1) or (3).
- However, \( A^k = 0 \) for \( k \geq 3 \) so \( x[3] = A^3 x[0] = 0 \) for any initial state \( x[0] \) and any input \( u[k] \).
- Thus, the system is *controllable to the origin* but not *controllable from the origin* or *reachable*.  

Definition (1) encompasses the other two definitions, so is used as our definition of controllable.

**Discrete-time observability**

- Can we reconstruct the state $x[0]$ from the output $y[k]$ and input $u[k]$?

  \[
  y[k] = Cx[k] + Du[k] \\
  y[0] = Cx[0] + Du[0] \\
  \vdots \\
  y[n-1] = C[A^{n-1}x[0] + A^{n-2}Bu[0] + \cdots + Bu[n-1]] + Du[n-1].
  \]

- In vector form, we can write

  \[
  \begin{bmatrix}
  y[0] \\
  y[1] \\
  \vdots \\
  y[n-1]
  \end{bmatrix} = \begin{bmatrix}
  C \\
  CA \\
  \vdots \\
  CA^{n-1}
  \end{bmatrix} x[0] + \begin{bmatrix}
  D & 0 & \cdots & 0 \\
  CB & D & \cdots & 0 \\
  CAB & CB & \cdots & D
  \end{bmatrix} \begin{bmatrix}
  u[0] \\
  u[1] \\
  \vdots \\
  u[n-1]
  \end{bmatrix}.
  \]

- So,

  \[
  x[0] = \mathcal{O}^{-1} \begin{bmatrix}
  y[0] \\
  \vdots \\
  y[n-1]
  \end{bmatrix} - \mathcal{T} \begin{bmatrix}
  u[0] \\
  \vdots \\
  u[n-1]
  \end{bmatrix}.
  \]

- If $\mathcal{O}$ is full-rank or nonsingular, $x[0]$ may be reconstructed with any $y[k], u[k]$. We say that \{C, A\} form an “observable pair.”
• Do more measurements of \( y[n], y[n+1], \ldots \) help in reconstructing \( x[0] \)? No! (Caley–Hamilton theorem: next section).

• So, if the original state is not “observable” with \( n \) measurements, then it will not be observable with more than \( n \) measurements either.

• Since we know \( u[k] \) and the dynamics of the system, if the system is observable we can determine the entire state sequence \( x[k], k \geq 0 \) once we determine \( x[0] \)

\[
x[n] = A^n x[0] + \sum_{i=0}^{n-1} A^{n-1-i} B u[k]
\]

\[
e = A^n O^{-1} \begin{bmatrix} y[0] \\ \vdots \\ y[n-1] \end{bmatrix} - T \begin{bmatrix} u[0] \\ \vdots \\ u[n-1] \end{bmatrix} + C \begin{bmatrix} u[n-1] \\ \vdots \\ u[0] \end{bmatrix}.
\]

• A perfectly good observer (no differentiators...)

5.4: Cayley–Hamilton Theorem

- A square matrix $A$ satisfies its own characteristic equation. That is, if
  \[ \chi(\lambda) = \det(\lambda I - A) = 0 \]
  then
  \[ \chi(A) = 0. \]
- We can easily show this if $A$ is diagonalizable. Let
  \[ A = V^{-1} \Lambda V. \]
- Then
  \[
  A^2 = V^{-1} \Lambda V V^{-1} \Lambda V \\
  = V^{-1} \Lambda^2 V \\
  A^k = V^{-1} \Lambda^k V.
  \]
- The characteristic polynomial is:
  \[ \chi(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1 \]
  so if we replace $\lambda$ with $A$ we get
  \[
  \chi(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1I \\
  = V^{-1} \left[ \Lambda^n + a_{n-1}\Lambda^{n-1} + \cdots + a_1I \right] V.
  \]
- To “prove” the Cayley–Hamilton theorem, we just need to show that
  the quantity inside the brackets is zero.
- It is a diagonal matrix, and each diagonal element has the form
  \[ \lambda_i^n + a_{n-1}\lambda_i^{n-1} + \cdots + a_1 = 0 \]
  because $\lambda_i$ is an eigenvalue of $A$. 
So, each diagonal element is zero, and we have shown the proof.

If \( A \) is not diagonalizable, the same proof may be repeated using the Jordan form and Jordan blocks:

\[
A = T^{-1}JT.
\]

Consider a sketch of the proof for a Jordan block of size 2 and

\[
\chi(\lambda_i) = \lambda_i^3 + a_2 \lambda_i^2 + a_1 \lambda_i + a_0 = 0.
\]

Then

\[
J_i = \begin{bmatrix}
\lambda_i & 1 \\
0 & \lambda_i
\end{bmatrix}
\]

\[
\chi(J_i) = \begin{bmatrix}
\lambda_i^3 & 3 \lambda_i^2 \\
0 & \lambda_i^3
\end{bmatrix} + a_2 \begin{bmatrix}
\lambda_i^2 & 2 \lambda_i \\
0 & \lambda_i^2
\end{bmatrix} + a_1 \begin{bmatrix}
\lambda_i & 1 \\
0 & \lambda_i
\end{bmatrix} + a_0 \begin{bmatrix}
1 & 0 \\
0 & 1
\end{bmatrix}.
\]

We can easily see that the diagonal and lower-diagonal components are zero so

\[
\chi(J_i) = \begin{bmatrix}
0 & \alpha \\
0 & 0
\end{bmatrix}
\]

where

\[
\alpha = 3 \lambda_i^2 + 2a_2 \lambda_i + a_1
\]

but \( \alpha = \frac{d}{d\lambda} \chi(\lambda) = 0 \) which completes the sketch.

**SIGNIFICANCE:** The Cayley–Hamilton theorem shows us that \( A^n \) is a function of matrix powers \( A^{n-1} \) down to \( A^0 \). Therefore, to compute any polynomial of \( A \) it suffices to compute only powers of \( A \) up to \( A^{n-1} \) and appropriately weight their sum. A lot of proofs use the Cayley–Hamilton theorem.
As we just saw with the section on discrete-time observability, the Cayley–Hamilton theorem implies that if we cannot observe the state with \( n \) measurements, we cannot observe it with more measurements either.

**EXAMPLE:** With \( A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \) we have \( \chi(\lambda) = \det(\lambda I - A) \), so

\[
\chi(\lambda) = \det \left( \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} - A \right)
\]

\[
\chi(\lambda) = \lambda^2 - 5\lambda - 2
\]

\[
\chi(A) = A^2 - 5A - 2I = \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} - 5 \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} - 2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\]

\[
= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.
\]

**Disclaimer: Does observability/controllability matter, practically?**

- The singularity of \( C \) has only one “bit” of information: Is the realization mathematically controllable or not? This may not tell the whole story.

\[
A = \begin{bmatrix} 1 & 0 \\ 0 & 1 + \epsilon \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \\ 1 & 1 + \epsilon \end{bmatrix}
\]

- \( \{A, B\} \) are a controllable pair, but barely.

**EXAMPLE:** Controlling an airplane. (Ideas only, no details). System state

\[
x \triangleq \begin{bmatrix} \theta & \dot{\theta} & \phi & \dot{\phi} \end{bmatrix}^T, \quad \theta = \text{Pitch}, \ \phi = \text{Roll}.
\]

- Control with elevator?
where $\delta_e$ is the elevator angle. $C$ is singular $\Rightarrow$ can’t influence roll with elevators.

- Control with ailerons?

$$\dot{x} = \begin{bmatrix} F_\theta & \epsilon \\ 0 & F_\phi \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \delta_a$$

where $\delta_a$ is the aileron angle. $C$ is nonsingular! So, we can control both pitch AND roll with ailerons.

- **THIS IS NONSENSE!** Physically think of the system! Do you want to roll plane over every time you need to pitch down?

- Physical intuition can be better than finding $C$. Other tools can help...
5.5: Continuous-time Gramians

Continuous-time controllability Gramian

- If a continuous-time system is controllable, then

\[ W_c(t) = \int_0^t e^{A \tau} BB^T e^{A^T \tau} \, d\tau \]

is nonsingular for \( t > 0 \).

**SIGNIFICANCE:** Consider

\[ x(t_1) = e^{At_1} x(0) + \int_0^{t_1} e^{A(t_1 - \tau)} Bu(\tau) \, d\tau. \]

- We claim that for any \( x(0) = x_0 \) and any \( x(t_1) = x_1 \) the input

\[ u(t) = -B^T e^{A^T(t_1 - t)} W_c^{-1}(t_1) [e^{At_1} x_0 - x_1] \]

will transfer \( x_0 \) to \( x_1 \) at time \( t_1 \).

**PROOF:** Substitute the expression for \( u(t) \) into the convolution expression:

\[
\begin{align*}
x(t_1) &= e^{At_1} x(0) - \int_0^{t_1} e^{A(t_1 - \tau)} BB^T e^{A^T(t_1 - \tau)} \, d\tau \ W_c^{-1}(t_1) [e^{At_1} x_0 - x_1] \\
&= e^{At_1} x(0) - \int_0^{t_1} e^{A \beta} BB^T e^{A^T \beta} \, d\beta \ W_c^{-1}(t_1) [e^{At_1} x_0 - x_1] \\
&= e^{At_1} x(0) - W_c(t_1) W_c^{-1}(t_1) [e^{At_1} x_0 - x_1] \\
&= e^{At_1} x(0) - e^{At_1} x_0 + x_1 = x_1.
\end{align*}
\]

- Therefore, we can **compute** the input \( u(t) \) required to transfer the state of the system from one state to another over an arbitrary interval of time. The solution is also the minimum-energy solution.
**EXAMPLE:** Consider the system in diagonal form

\[
\dot{x}(t) = \begin{bmatrix} -0.5 & 0 \\ 0 & -1 \end{bmatrix} x(t) + \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} u(t).
\]

- The controllability matrix is:

\[
C = \begin{bmatrix} 0.5 & -0.25 \\ 1 & -1 \end{bmatrix}
\]

which has rank 2, so the system is controllable.

- Consider the input required to move the system state from \(x(0) = [10 \ -1]^T\) to zero in two seconds. 

\[
W_c(2) = \int_0^2 \left( \begin{bmatrix} e^{-0.5\tau} & 0 \\ 0 & e^{-\tau} \end{bmatrix} \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} \begin{bmatrix} e^{-0.5\tau} & 0 \\ 0 & e^{-\tau} \end{bmatrix} \right) d\tau
\]

\[
= \begin{bmatrix} 0.2162 & 0.3167 \\ 0.3167 & 0.4908 \end{bmatrix},
\]

and

\[
u(t) = -\begin{bmatrix} 0.5 & 1 \end{bmatrix} \begin{bmatrix} e^{-0.5(2-t)} & 0 \\ 0 & e^{-(2-t)} \end{bmatrix} W_c(2)^{-1} \begin{bmatrix} e^{-1} & 0 \\ 0 & e^{-2} \end{bmatrix} \begin{bmatrix} 10 \\ -1 \end{bmatrix}
\]

\[
= -58.82 e^{0.5t} + 27.96 e^t.
\]

- If a continuous-time system is controllable, and if it is also stable, then

\[
W_c = \int_0^\infty e^{At} BB^T e^{A^T\tau} d\tau
\]

can be found by solving for the unique (positive-definite) solution to the (Lyapunov) equation

\[
AW_c + W_c A^T = -BB^T.
\]

\(W_c\) is called the *controllability Gramian.*
**PROOF:** We proved this identity when considering Lyapunov stability.

- $W_c$ measures the minimum energy required to reach a desired point $x_1$ starting at $x(0) = 0$ (with no limit on $t$)

\[\min \left\{ \int_0^t \|u(\tau)\|^2 \, d\tau \mid x(0) = 0, \, x(t) = x_1 \right\} = x_1^T W_c^{-1} x_1.\]

- In fact, for any specific “$t$”, the minimum energy is $x_1^T W_c^{-1}(t)x_1$.

- If $A$ is stable, $W_c^{-1} > 0$ which implies “we can’t get anywhere for free”.

- If $A$ is unstable, then $W_c^{-1}$ can have a nonzero nullspace $W_c^{-1}z = 0$ for some $z \neq 0$ which means that we can get to $z$ using $u$’s with energy as small as you like! ($u$ just gives a little kick to the state; the instability carries it out to $z$ efficiently).

- $W_c$ may be a better indicator of controllability than $C$.

**Continuous-time observability Gramian**

- If a system is observable, $W_o(t)$ is nonsingular for $t > 0$ where

\[W_o(t) = \int_0^t e^{A^T \tau} C^T C e^{A \tau} \, d\tau.\]

**SIGNIFICANCE:** We can prove that

\[x(0) = W_o^{-1}(t_1) \int_0^{t_1} e^{A^T \tau} C^T \tilde{y}(\tau) \, d\tau\]

where

\[\tilde{y}(\tau) = y(\tau) - C \int_0^\tau e^{A(t-\tau)} Bu(\tau) \, d\tau - Du(\tau) = C e^{A t} x(0)\]

- Therefore, we can determine the initial state $x(0)$ given a finite observation period (and not use differentiators!).
PROOF: We prove that the above equations are correct by substitution:

\[ W_o^{-1}(t_1) \int_0^{t_1} e^{AT} C^T y(t) \, dt = W_o^{-1}(t_1) \int_0^{t_1} e^{AT} C C e^{At} \, dt \, x(0) \]
\[ = W_o^{-1}(t_1) W_o(t_1) x(0) \]
\[ = x(0). \]

- If a continuous-time system is observable, and if it is also stable, then
  \[ W_o = \int_0^{\infty} e^{AT} C^T C e^{At} \, d\tau \]
  is the unique (positive-definite) solution to the (Lyapunov) equation
  \[ A^T W_o + W_o A = -C^T C. \]

  \( W_o \) is called the observability Gramian.

- This relationship can be proven in the same way we proved the similar relationship for the controllability gramian.

- If measurement (sensor) noise is \( \text{IID } \mathcal{N}(0, \sigma^2 I) \) then \( W_o \) is a measure of error covariance in measuring \( x(0) \) from \( u \) and \( y \) over longer and longer periods

  \[ \lim_{t \to \infty} \mathbb{E} \left\| \hat{x}(0) - x(0) \right\|^2 = \sigma x(0)^T W_o^{-1} x(0). \]

- If \( A \) is stable, then \( W_o^{-1} > 0 \) and we can’t estimate the initial state perfectly even with an infinite number of measurements \( u(t) \) and \( y(t) \) for \( t \geq 0 \) (since memory of \( x(0) \) fades).

- If \( A \) is not stable then \( W_o^{-1} \) can have a nonzero nullspace \( W_o^{-1} x(0) = 0 \) which means that the covariance goes to zero as \( t \to \infty \).

- \( W_o \) may be a better indicator of observability than \( O \).
5.6: Discrete-time Gramians

Discrete-time controllability Gramian

- In discrete-time, if a system is controllable, then
  \[ W_{dc}[n - 1] = \sum_{m=0}^{n-1} A^m BB^T (A^T)^m \]

  is nonsingular. In particular,
  \[ W_{dc} = \sum_{m=0}^{\infty} A^m BB^T (A^T)^m \]

  is called the discrete-time controllability Gramian and is the unique positive-definite solution to the Lyapunov equation
  \[ W_{dc} - AW_{dc}A^T = BB^T. \]

- As with continuous-time, \( W_{dc} \) measures the minimum energy required to reach a desired point \( x_1 \) starting at \( x[0] = 0 \) (with no limit on \( m \))
  \[ \min \left\{ \sum_{k=0}^{m} \|u[k]\|^2 \mid x[0] = 0, \ x[m] = x_1 \right\} = x_1^T W_{dc}^{-1} x_1. \]

  **ASIDE:** When considering discrete-time stability, we showed that this form of equation is indeed a Lyapunov equation.

Discrete-time observability Gramian

- In discrete-time, if a system is observable, then
  \[ W_{do}[n - 1] = \sum_{m=0}^{n-1} (A^T)^m CC^T A^m \]

  is nonsingular. In particular,
\[ W_{do} = \sum_{m=0}^{\infty} (A^T)^m C C^T A^m \]

is called the \textit{discrete-time observability Gramian} and is the unique positive-definite solution to the Lyapunov equation

\[ W_{do} - A^T W_{do} A = C^T C. \]

- As with continuous-time, if measurement (sensor) noise is IID \( \mathcal{N}(0, \sigma I) \) then \( W_{do} \) is a measure of error covariance in measuring \( x[0] \) from \( u \) and \( y \) over longer and longer periods

\[
\lim_{t \to \infty} \mathbb{E} \| \hat{x}[0] - x[0] \|^2 = \sigma x[0]^T W_{do}^{-1} x[0].
\]
5.7: Computing transformation matrices

Transformation to controllability form

- Given a system

\[
\begin{align*}
\dot{x}(t) &= Ax(t) + Bu(t) \\
y(t) &= Cx(t) + Du(t),
\end{align*}
\]

can we find a transformation (similarity) matrix \( T \) to transform this system into controllability form? Recall, this looks like:

\[
\begin{align*}
\dot{x}_{co}(t) &= \begin{bmatrix}
0 & \cdots & 0 & -a_n \\
1 & \vdots & -a_{n-1} \\
\vdots & \ddots & \vdots \\
0 & \cdots & 1 & -a_1
\end{bmatrix} x_{co}(t) + \begin{bmatrix}
1 \\
0 \\
\vdots \\
0
\end{bmatrix} u(t) \\
y(t) &= \begin{bmatrix}
\beta_1 & \beta_2 & \cdots & \beta_n
\end{bmatrix} x_{co}(t) + Du(t).
\end{align*}
\]

- Note that \( C_{co} = I \) so it is controllable. Thus, our original system must be controllable.

- The transformation is accomplished via

\[ x = T x_{co}. \]

- Thus

\[
\begin{align*}
T^{-1} AT &= A_{co}, \\
T^{-1} B &= B_{co} \\
CT &= C_{co}, \\
D &= D_{co}.
\end{align*}
\]

- Let’s find \( T \) explicitly; let

\[ T = [t_1, \ldots, t_n]. \]
Note that $B = TB_{co}$ or

$$B = T \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = t_1.$$ 

So, $t_1 = B$. From $AT = TA_{co}$ we have

$$A[t_1, \ldots , t_n] = [t_1, \ldots , t_n] \begin{bmatrix} 0 & \cdots & 0 & -a_n \\ 1 & \vdots & -a_{n-1} \\ \vdots & \ddots & \ddots \\ 0 & \cdots & 1 & -a_1 \end{bmatrix}.$$ 

So

$$[At_1, \ldots At_n] = \begin{bmatrix} t_2, \ldots , t_n, -\sum_{k=1}^{n} a_k t_{n-k+1} \end{bmatrix}.$$ 

By induction,

$$At_1 = t_2 = AB$$

$$At_2 = t_3 = A^2 B,$$ and so forth...

so

$$T = [B \ AB \ \cdots \ A^{n-1} B] = C.$$ 

**CONCLUSION:** A system $\{A, B, C, D\}$ can be transformed to controllability canonical form if and only if it is controllable, in which case the change of coordinates is

$$x = Cx_{co}.$$ 

**EXTENSION I:** If $x_{old} = Tx_{new}$ then $T = C_{old}C_{new}^{-1}$. That is, to convert between any two realizations, $T$ is a combination of the controllability
matrices of the two different realizations.

\[
\begin{align*}
C_{\text{new}} &= \begin{bmatrix} B_{\text{new}} & A_{\text{new}} B_{\text{new}} & \cdots & A_{\text{new}}^{n-1} B_{\text{new}} \end{bmatrix} \\
&= \begin{bmatrix} T^{-1} B_{\text{old}} & T^{-1} A_{\text{old}} T T^{-1} B_{\text{old}} & \cdots & T^{-1} A_{\text{old}}^{n-1} T T^{-1} B_{\text{old}} \end{bmatrix} \\
&= T^{-1} C_{\text{old}},
\end{align*}
\]

or

\[
T = C_{\text{old}} C_{\text{new}}^{-1}.
\]

**EXTENSION II:** If \( x_{\text{old}} = T x_{\text{new}} \) then \( T = O_{\text{old}}^{-1} O_{\text{new}} \). This can be shown in a similar way.
5.8: Canonical (Kalman) decompositions

- What happens if \{A, B\} not controllable or if \{A, C\} not observable?
  - Is “part” of the system controllable?
  - Is “part” of the system observable?
- Given a system with \{A, B\} not controllable, \( \dot{x}(t) = Ax(t) + Bu(t) \), \( y(t) = Cx(t) \), let’s try to transform to controllability canonical form.
- Let \( t_1 = B, t_2 = AB, \ldots t_r = A^{r-1}B \), and suppose \( t_1, t_2, \ldots, t_r \) are independent but
  \[
  t_{r+1} = A^r B = -\alpha_r t_1 - \cdots - \alpha_1 t_r
  \]
  for some constants \( \alpha_1, \ldots, \alpha_r \).
- Then \( \text{rank}(C) = r \) since the vectors \( A^r B, A^{r+1} B, \ldots, A^{n-1} B \) can all be expressed as a linear combination of \( t_1, t_2, \ldots, t_r \).
- Let \( s_{r+1}, \ldots, s_n \) be your favorite vectors for which
  \[
  \bar{C}_{\text{old}} = \begin{bmatrix} t_1 & \cdots & t_r & s_{r+1} & \cdots & s_n \end{bmatrix}
  \]
  is invertible.
  - If we were able to change coordinates to controllability form via
    \( x_{\text{old}} = Tx_{\text{new}} \), we would use \( T = C_{\text{old}}C_{\text{new}}^{-1} \).
  - But, the system is not controllable, so we use \( T = \bar{C}_{\text{old}}C_{\text{new}}^{-1} = \bar{C}_{\text{old}} \).
- We get (in the new coordinate system)
  \[
  \begin{bmatrix}
  \dot{x}_c(t) \\
  \dot{x}_{\bar{c}}(t)
  \end{bmatrix}
  = \begin{bmatrix}
  A_c & A_{12} \\
  0 & A_{\bar{c}}
  \end{bmatrix}
  \begin{bmatrix}
  x_c(t) \\
  x_{\bar{c}}(t)
  \end{bmatrix}
  + \begin{bmatrix}
  B_c \\
  0
  \end{bmatrix}
  u(t)
  \]
  \[
  y(t) = \begin{bmatrix}
  C_c & C_{\bar{c}}
  \end{bmatrix}
  \begin{bmatrix}
  x_c(t) \\
  x_{\bar{c}}(t)
  \end{bmatrix}
  + Du(t).
  \]
- $A_c$ is a right-companion matrix and $B_c$ is of the controllability-canonical form.

- We see that the uncontrollable modes $x_c$ are completely decoupled from $u(t)$.

**EXAMPLE:** Consider

\[
\frac{1}{s + 1} = \frac{s - 1}{(s + 1)(s - 1)} = \frac{s - 1}{s^2 - 1}.
\]

- In observer-canonical form,

\[
\begin{align*}
\dot{x}(t) &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ -1 \end{bmatrix} u(t) \\
y(t) &= \begin{bmatrix} 1 & 0 \end{bmatrix} x(t).
\end{align*}
\]

- So,

\[
t_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{and} \quad t_2 = At_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -t_1.
\]

- So, $\text{rank}(C) = 1$. Let $s_1 = [1 \ 0]^T$. The converted state-space form is

\[
\begin{align*}
\dot{x}(t) &= \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix} \tilde{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t) \\
y(t) &= \begin{bmatrix} 1 & 1 \end{bmatrix} \tilde{x}(t).
\end{align*}
\]

- Now suppose that we desire to transform a system that is not observable. The dual form separates out the unobservable states.

- Let $t_1 = C$, $t_2 = CA$, \ldots $t_r = CA^{r-1}$, and suppose $t_1, t_2, \ldots, t_r$ are independent but

\[
t_{r+1} = CA^r = -\alpha_r t_1 - \cdots - \alpha_1 t_r
\]

for some constants $\alpha_1, \ldots, \alpha_r$. 

Then, let \( s_{r+1}, \ldots, s_n \) be your favorite row vectors for which

\[
\tilde{O}_{\text{old}} = \begin{bmatrix}
    t_1 \\
    \vdots \\
    t_r \\
    \cdots \\
    s_{r+1} \\
    \vdots \\
    s_n
\end{bmatrix}
\]

is invertible.

- If we were able to change coordinates to observability form via \( x_{\text{old}} = T x_{\text{new}} \), we would use \( T = O^{-1}_{\text{old}} O_{\text{new}} \).
- But, the system is not observable, so we use \( T = \tilde{O}^{-1}_{\text{old}} O_{\text{new}} = \tilde{O}^{-1}_{\text{old}} \).

Then, we get the Kalman decomposition

\[
\begin{bmatrix}
    \dot{x}_o(t) \\
    \dot{x}_{\bar{o}}(t)
\end{bmatrix} = \begin{bmatrix}
    A_o & 0 \\
    A_{21} & A_{\bar{o}}
\end{bmatrix} \begin{bmatrix}
    x_o(t) \\
    x_{\bar{o}}(t)
\end{bmatrix} + \begin{bmatrix}
    B_o \\
    B_{\bar{o}}
\end{bmatrix} u(t)
\]

\[
y(t) = \begin{bmatrix}
    C_o & 0
\end{bmatrix} \begin{bmatrix}
    x_o(t) \\
    x_{\bar{o}}(t)
\end{bmatrix} + Du(t).
\]

Note: No path from \( x_{\bar{o}} \) to \( y \)!

**Full Kalman decomposition**

- We can do both transformations in sequence to get full Kalman decomposition. See text chapter 16.
5.9: Popov–Belevitch–Hautus controllability/observability tests

**PBH EIGENVECTOR TEST:** \( \{C, A\} \) is an unobservable pair iff a non-zero eigenvector \( v \) of \( A \) satisfies \( C v = 0 \). *(i.e., \( C \) and \( v \) are perpendicular).*

**PROOF:** \( \Rightarrow \) Suppose \( A v = \lambda v \) and \( C v = 0 \) for \( v \neq 0 \). Then \( C A v = \lambda C v = 0 \) and so forth up to \( C A^{n-1} v = \lambda^{n-1} C v = 0 \). So, \( O v = 0 \) and since \( v \neq 0 \) this means that \( O \) is not full rank and \( \{C, A\} \) is not observable.

\( \Leftarrow \) Now suppose that \( O \) is not full rank (\( \{C, A\} \) unobservable). Let’s extract the unobservable part. That is, find \( T \) such that

\[
T^{-1} A T = \begin{bmatrix}
A_o & 0 \\
A_{21} & A_{\bar{o}}
\end{bmatrix}
\]

\[
C T = \begin{bmatrix}
C_o & 0
\end{bmatrix},
\]

where \( A_o \) is size \( r \) where \( r = \text{rank}(O) \) and therefore \( A_{\bar{o}} \) is size \( n - r \).

- Let \( v_2 \neq 0 \) be an eigenvector of \( A_{\bar{o}} \). Then

\[
T^{-1} A T \begin{bmatrix}
0 \\
v_2
\end{bmatrix} = \begin{bmatrix}
A_o & 0 \\
A_{21} & A_{\bar{o}}
\end{bmatrix} \begin{bmatrix}
0 \\
v_2
\end{bmatrix} = \lambda \begin{bmatrix}
0 \\
v_2
\end{bmatrix}
\]

so we have found an eigenvector \( z \) of \( A \)

\[
A z = \lambda z \quad \text{where} \quad z = T \begin{bmatrix}
0 \\
v_2
\end{bmatrix}.
\]

- Now, we just need to show that \( C z = 0 \) (note: \( z \neq 0 \)).

\[
C z = \overbrace{C T}^{\text{C}} \begin{bmatrix}
0 \\
v_2
\end{bmatrix} = \begin{bmatrix}
C_o & 0
\end{bmatrix} \begin{bmatrix}
0 \\
v_2
\end{bmatrix} = 0
\]

and we are done.
**DUAL:** \( \{A, B\} \) is uncontrollable iff there is a left eigenvector \( w^T \) of \( A \) such that \( w^T B = 0 \).

**INTERPRETATION:** In modal coordinates, homogeneous response is

\[
x(t) = \sum_{i=1}^{n} e^{\lambda_i t} v_i(w_i^T x(0)) \quad \text{and} \quad y(t) = C x(t).
\]

Or,

\[
y(t) = \sum_{i=1}^{n} e^{\lambda_i t} C v_i(w_i^T x(0)).
\]

If \( \{C, A\} \) is unobservable, then it has an *unobservable mode*, where \( A v_i = \lambda_i v_i \) and \( C v_i = 0 \).

- If \( \{A, B\} \) is uncontrollable, then it has an *uncontrollable mode*, namely the coefficients of the state along that mode is independent of the input \( u(t) \).
  - The coefficients of \( x \) in the mode associated with \( \lambda \) are \( w^T x \).
    \[
    \frac{d}{dt}(w^T x) = w^T (Ax + Bu) = \lambda (w^T x)
    \]
    or
    \[
    w^T x(t) = e^{\lambda t} (w^T x(0))
    \]
    regardless of the input \( u(t) \)!

**PBH RANK TESTS:** The following two tests are often easier to perform

1. \( \{A, B\} \) controllable iff \( \text{rank} \left[ \begin{array}{c|c} sI - A & B \end{array} \right] = n \) for all \( s \in \mathbb{C} \).
   \[
   \Rightarrow \text{If} \text{ rank} \left[ \begin{array}{c|c} sI - A & B \end{array} \right] = n \text{ for all } s \in \mathbb{C} \text{ then there can be no nonzero vector } v \text{ such that } v^T \left[ \begin{array}{c|c} sI - A & B \end{array} \right] = \left[ v^T (sI - A) \mid v^T B \right] = 0.
   \]
Consequently, there is no nonzero vector \( v \) such that \( v^T s = v^T A \) and \( v^T B = 0 \). By the PBH eigenvector test, the system will therefore be controllable.

\[ \Leftrightarrow \text{If the system is controllable, then there is no nonzero vector} \ v \ \text{such that} \ v^T s = v^T A \ \text{and} \ v^T B = 0 \ \text{by the PBH eigenvector test.} \]

Therefore, \( \text{rank} \begin{bmatrix} sI - A & B \end{bmatrix} = n \) for all \( s \in \mathbb{C} \).

2. \( \{C, A\} \) observable iff \( \text{rank} \begin{bmatrix} C \\ sI - A \end{bmatrix} = n \) for all \( s \in \mathbb{C} \). Proof similar.

**COMMENTS:** \( \text{rank} \begin{bmatrix} sI - A & B \end{bmatrix} = n \) for all \( s \) not eigenvalues of \( A \), so the test is really \( \{A, B\} \) controllable iff \( \text{rank} \begin{bmatrix} \lambda_i I - A & B \end{bmatrix} = n \) for \( \lambda_i, i = 1, \ldots, n \) the eigenvalues of \( A \). (Dual argument for observability).

- If \( \begin{bmatrix} sI - A & B \end{bmatrix} \) drops rank at \( s = \lambda \) then there is an uncontrollable mode with exponent (frequency) \( \lambda \).
- If \( \begin{bmatrix} C \\ sI - A \end{bmatrix} \) drops rank at \( s = \lambda \) then there is an unobservable mode with exponent (frequency) \( \lambda \).

**Summary**

- Therefore, we can label *individual modes* of a system as either controllable or not, or observable or not.
- The overall picture is:
- Some other definitions:

**STABILIZABLE:** A system whose unstable modes are controllable.

**DETECTABLE:** A system whose unstable modes are observable.
5.10: Minimal realizations—Why a system isn’t control/observable

- To realize \( H(s) = \frac{1}{s+1} \) we could use

1. \( \dot{x}(t) = -x(t) + u(t), \ y(t) = x(t) \). This gives
   \[
   A = [-1], \ B = [1], \ C = [1], \ D = [0].
   \]
   \[
   \frac{C \ \text{adj}(sI - A)B}{\text{det}(sI - A)} = \frac{1}{s+1}.
   \]
   This realization is both controllable and observable.

2. Observer realization of
   \[
   \frac{1}{s+1} \frac{s-1}{s^2-1} = \frac{s-1}{s^2-1}.
   \]
   This gives
   \[
   A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ B = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ C = \begin{bmatrix} 1 & 0 \end{bmatrix}, \ D = [0].
   \]
   \[
   \frac{C \ \text{adj}(sI - A)B}{\text{det}(sI - A)} = \frac{s-1}{s^2-1} = \frac{1}{s+1}.
   \]
   \[
   \mathcal{O} = \begin{bmatrix} C \\ CA \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \text{ Observable.}
   \]
   \[
   C = \begin{bmatrix} B & AB \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \text{ Not controllable.}
   \]

3. Controller realization of
   \[
   \frac{1}{s+1} \frac{s+10}{s+10} = \frac{s+10}{s^2 + 11s + 10}.
   \]
   This gives
   \[
   A = \begin{bmatrix} -11 & -10 \\ 1 & 0 \end{bmatrix}, \ B = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ C = \begin{bmatrix} 1 & 10 \end{bmatrix}, \ D = [0].
   \]
   \[
   \frac{C \ \text{adj}(sI - A)B}{\text{det}(sI - A)} = \frac{s+10}{s^2 + 11s + 10} = \frac{1}{s+1}.
   \]
   \[
   C = \begin{bmatrix} B & AB \end{bmatrix} = \begin{bmatrix} 1 & -11 \\ 0 & 1 \end{bmatrix}. \text{ Controllable.}
   \]
\[ O = \begin{bmatrix} C \\ CA \end{bmatrix} = \begin{bmatrix} 1 & 10 \\ -1 & -10 \end{bmatrix}. \] Not observable.

**TREND:** Non-minimal realizations of transfer functions will either be uncontrollable, unobservable, or both.

- Four equivalent statements:
  
  **I:** There exist common roots of \( C \ adj(sI - A) B \) and \( \det(sI - A) \).
  
  **II:** There exist eigenvalues of \( A \) which are not poles of \( G(s) \), counting multiplicities.
  
  **III:** The system is either unobservable or uncontrollable.
  
  **IV:** There exist extra (unnecessary) states—non minimal.

**DEFINITION:** We say a system is minimal if no system with the same transfer function has fewer states.

**PROOF:** \( I \iff II \)

**I \iff II:** The transfer function

\[
G(s) = C(sI - A)^{-1}B + D = \frac{C \ adj(sI - A) B + D \ det(sI - A)}{\det(sI - A)}.
\]

If there are common roots in \( C \ adj(sI - A) B \) and \( \det(sI - A) \) they will cancel out of the transfer function. But, eigenvalues of \( A \) are \( \det(sI - A) = 0 \), so poles of \( G(s) \) will not contain all eigenvalues of \( A \).

**II \iff I:** Eigenvalues of \( A \) are \( \det(sI - A) = 0 \). Poles of \( G(s) \), from above, are \( \det(sI - A) = 0 \) unless canceled. The only way to cancel a pole is to have common root in \( C \ adj(sI - A) B \).
PROOF: \( I \iff IV \) \( \{A, B, C, D\} \) is minimal iff \( C \ adj(sI - A)B \) and \( \det(sI - A) \) are coprime (have no common roots).

\( I \iff IV \): Suppose \( C \ adj(sI - A)B \) and \( \det(sI - A) \) have common roots.

- Cancel to get

\[
G(s) = \frac{C \ adj(sI - A)B}{\det(sI - A)} + D = \frac{b_r(s)}{a_r(s)}
\]

where \( b_r(s) \) and \( a_r(s) \) are coprime ("r" means "reduced").
- Because of cancellation, \( k = \deg(a_r) < \deg(\det(sI - A)) = n \).
- Consider controller canonical form realization of \( b_r(s)/a_r(s) \), for example.
- It has \( k \) states, but same transfer function as \( \{A, B, C, D\} \), contradicting that \( \{A, B, C, D\} \) minimal.

\( IV \iff I \): If there are extra states (non-minimal) then \( n > k \) (\( n = \# \) states, \( k = \# \) poles in \( G(s) \)).

\[
G(s) = C(sI - A)^{-1}B + D = \frac{C \ adj(sI - A)B + D \ det(sI - A)}{\det(sI - A)}
\]

\( G(s) \) has \( n \) poles unless some cancel with \( C \ adj(sI - A)B \). Therefore, if \( n > k \), \( C \ adj(sI - A)B \) and \( \det(sI - A) \) are not coprime.

PROOF: \( III \iff IV \) Controllable and observable iff minimal.

\( III \iff IV \): Uncontrollable or unobservable \( \implies \) not minimal.

Perform Kalman decomposition to split system into \( co, c\bar{o}, \bar{c}o \) and \( \bar{c}\bar{o} \) parts.
\[
\tilde{A} = \begin{bmatrix}
A_c & A_{12} \\
0 & A_{\tilde{c}}
\end{bmatrix}, \quad \tilde{B} = \begin{bmatrix}
B_c \\
0
\end{bmatrix}, \quad \tilde{C} = \begin{bmatrix}
C_c & C_{\tilde{c}}
\end{bmatrix}, \quad \tilde{D} = [0].
\]

Same transfer function using \(A_c, B_c, C_c\) as \(\tilde{A}, \tilde{B}, \tilde{C}\). Therefore uncontrollable and/or unobservable means not minimal.

**IV \implies III:** Non-minimal means uncontrollable or unobservable.

- Suppose \(\{A, B, C, D\}\) is non-minimal.

\[
\frac{C(sI - A)^{-1}B + D}{n \text{ states}} = \frac{\tilde{C}(sI - \tilde{A})^{-1}\tilde{B} + \tilde{D}}{r < n \text{ states}}
\]

\[
\frac{CB}{s} + \frac{CAB}{s^2} + \frac{CA^2B}{s^3} + \cdots = \frac{\tilde{C} \tilde{B}}{s} + \frac{\tilde{C} \tilde{A} \tilde{B}}{s^2} + \frac{\tilde{C} \tilde{A}^2 \tilde{B}}{s^3} + \cdots
\]

- Consider

\[
\mathcal{O}C = \begin{bmatrix}
C \\
CA \\
CA^2 \\
\vdots \\
CA^{n-1}
\end{bmatrix} \begin{bmatrix}
B & AB & A^2B & \cdots & A^{n-1}B
\end{bmatrix}
\]
FACT: All minimal realization of $G(s)$ are related by a unique change of coordinates $T$. Can you prove this?

Where to from here?

- Have seen important topics of observability and controllability.
- Now understand how a system could be unobservable or uncontrollable.
  - If physically true, may need to add sensors or actuators.
- Important to work with minimal realizations, when possible.
- Now, time to consider: How do I build a controller?
  - Very powerful tools in next chapter.

Therefore $\det(\mathcal{O}) \det(C) = 0$, so the system is either unobservable, or uncontrollable, or both.

The four equivalences have now been proven.