Module 1-3: Controllability and Observability

Linear Control Systems (2020)

Ding Zhao

Assistant Professor

College of Engineering

School of Computer Science

Carnegie Mellon University

Table of Contents

1. Controllability & Observability Matrices
   - Solutions of Simultaneous Linear Equations
2. Popov-Belevitch-Hautus Tests
3. Jordan Form Tests
   - Singular Value Decomposition
   - Gram-Schmidt Process
   - QR Decomposition

Controllability & Observability

Kalman, 1930-2016

- **Controllability**:
  A system is controllable if $\exists u(t), t \in [t_0, t_1]$ that transfers the system from any $x(t_0)$ to any $x(t_1)$.
  Heuristically, can we influence all the states (differently).
- **Observability**:
  A system is observable if knowing $u(t), y(t), t \in [t_0, t_1]$ is sufficient to uniquely solve for $\forall x(t_0)$.
  Heuristically, can we infer all internal states of a system from the input and output.

Table of Contents

1 Controllability & Observability Matrices
- Solutions of Simultaneous Linear Equations

2 Popov-Belevitch-Hautus Tests

3 Jordan Form Tests
- Singular Value Decomposition
- Gram-Schmidt Process
- QR Decomposition

Controllability for DT LTI Systems

We start looking at these for DT systems. The solution of the LTI discrete time system is:

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

$x[k] \in \mathbb{R}^{n\times 1}, A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, u[k] \in \mathbb{R}^{m\times 1}$.

$$\Rightarrow x[k] - A^k x[0] = [B \vdots AB \vdots \cdots \vdots A^{k-1}B] \begin{bmatrix} u[k-1] \\ u[k-2] \\ \vdots \\ u[0] \end{bmatrix} = \hat{P} \begin{bmatrix} u[k-1] \\ u[k-2] \\ \vdots \\ u[0] \end{bmatrix} = \hat{P} u$$

Let $x[k] - A^k x[0] = z \Rightarrow \hat{P}u = z, z \in \mathbb{R}^{n\times 1}, u \in \mathbb{R}^{km\times 1}, \hat{P} \in \mathbb{R}^{n\times km}$. We need to make sure "simultaneous linear equation" $\hat{P}u = z$ always have a solution. Fortunately, we have a theorem on it.

Solutions of Simultaneous Linear Equations

Consider

$$\mathbf{A}\mathbf{x} = \mathbf{y}$$

where $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$ and $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_n] : \mathbb{R}^n \to \mathbb{R}^m$

$$\mathbf{y} = \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \cdots + \mathbf{a}_n x_n, \mathbf{W} = [\mathbf{A} \vdots \mathbf{y}]$$

- **A solution exists:**
  iff $y \in \mathscr{R}(\mathbf{A}) \Leftrightarrow r(\mathbf{A}) = r(\mathbf{W}) \Leftrightarrow y$ is linearly dependent on columns of $\mathbf{A}$.
- **A solution does not exist:**
  iff $y \notin \mathscr{R}(\mathbf{A}) \Leftrightarrow r(\mathbf{A}) < r(\mathbf{W}) \Leftrightarrow y$ is linearly independent on columns of $\mathbf{A}$.
- **A unique solution exists:**
  iff $r(\mathbf{A}) = r(\mathbf{W}) = n \Leftrightarrow y$ is linearly dependent on columns of $\mathbf{A}$ and columns of $\mathbf{A}$ are independent
- **Multiple (actually infinite) solutions:**
  iff $r(\mathbf{A}) = r(\mathbf{W}) < n \Leftrightarrow y$ is linearly dependent on columns of $\mathbf{A}$ and columns of $\mathbf{A}$ are dependent

Multiple Solutions Case

$$\mathbf{A}\mathbf{x} = \mathbf{y}$$

$r(\mathbf{A}) = r([\mathbf{A} \vdots \mathbf{y}]) < n$

General solutions:

$$\mathbf{x} = \mathbf{x}_p + \alpha \mathbf{x}_n$$

where $\mathbf{x}_n \in \mathscr{N}(\mathbf{A})$, i.e. $\mathbf{A}\mathbf{x}_n = \mathbf{0}$ and $\mathbf{x}_n \neq \mathbf{0}$

$\mathbf{x}_p$ is a particular solution of $\mathbf{A}\mathbf{x} = \mathbf{y}$, i.e. $\mathbf{A}\mathbf{x}_p = \mathbf{y}$

α is an arbitrary scalar.

Check Rank with Gaussian Elimination

Convert the matrix $[\mathbf{A} \vdots \mathbf{y}]$ to echelon form using Gaussian Elimination

1. Convert to an upper triangular matrix.
2. Multiply rows by scalars, interchange rows, and/or add multiples of rows together.
3. Rank is the the number of nonzero rows

Example: Overactuated System

$\hat{P}u = z, \hat{P} \in \mathbb{R}^{n \times km}, u \in \mathbb{R}^{km \times 1}, z \in \mathbb{R}^{n \times 1}$. If $n < km$ (a common situation), we will have an **overactuated system**. Example:

$$\begin{bmatrix} 1 & -1 & 2 \\ -1 & 2 & 0 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 8 \\ 2 \end{bmatrix}$$

$$[\mathbf{A} \vdots \mathbf{y}] = \begin{bmatrix} 1 & -1 & 2 & 8 \\ -1 & 2 & 0 & 2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -1 & 2 & 8 \\ 0 & 1 & 2 & 10 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 4 & 18 \\ 0 & 1 & 2 & 10 \end{bmatrix}$$

x_3 is the "free" variable (no pivot in the third column).
Let $x_3 = 1$ and solve for x_1, x_2 to find the null space.

$$\begin{bmatrix} 1 & 0 & 4 & 0 \\ 0 & 1 & 2 & 0 \end{bmatrix}$$

$x_1 + 4 = 0, x_2 + 2 = 0, \Rightarrow x_1 = -4, x_2 = -2$

Example: Underactuated System Cont.

$$\begin{bmatrix} -4 \\ -2 \\ 1 \end{bmatrix} \text{ is a basis for the null space}$$

Particular Solution: Let $x_3 = 0$ [Since it is free variable, it doesn't change the solution.]

$$x_1 = 18, x_2 = 10$$

$$\mathbf{x} = \begin{bmatrix} 18 \\ 10 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -4 \\ -2 \\ 1 \end{bmatrix}$$

Controllability for LTI Systems

We start looking at these for DT systems. The solution of the LTI discrete time system is:

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

$$\Rightarrow x[k] - A^k x[0] = [B \vdots AB \vdots \cdots \vdots A^{k-1}B] \begin{bmatrix} u[k-1] \\ u[k-2] \\ \vdots \\ u[0] \end{bmatrix} = \hat{P} \begin{bmatrix} u[k-1] \\ u[k-2] \\ \vdots \\ u[0] \end{bmatrix} = \hat{P}U$$

$\Rightarrow$ To reach any state, $\hat{P} = [B \vdots AB \vdots \cdots \vdots A^{k-1}B] \in \mathbb{R}^{n \times km}$ must have rank n (for large k).

Let $P = [B \vdots AB \vdots \cdots \vdots A^{n-1}B] \in \mathbb{R}^{n \times n}$. Cayley-Hamilton Theorem $\Rightarrow$

$$\operatorname{rank}(P) = \operatorname{rank}(\hat{P})$$

Test Controllability

Test controllability using $\mathrm{rank}(P)$

A DT LTI system is controllable $\Leftrightarrow \mathrm{rank}(P) = n$, where

$$P = [B \vdots AB \vdots \cdots \vdots A^{n-1}B]$$

Test Observability

We can similarly extend our results to observability.

Test observability using $\mathrm{rank}(Q)$

A DT LTI system is observable $\Leftrightarrow \mathrm{rank}(Q) = n$, where

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

Note: Controllability and observability are dual aspects of the same problem, e.g. we can test the observability of a pair (A, C) by using the controllability tests on the pair (A^T, C^T).

Proof

$y(k) = CA^k x(0) + \sum_{m=0}^{k-1} CA^{k-m-1} Bu(m) + Du(k)$
Let $w[k] = y(k) - \sum_{m=0}^{k-1} CA^{k-m-1} Bu(m) - Du(k) = CA^k x[0]$

$$\begin{bmatrix} w[0] \\ w[1] \\ \vdots \\ w[k-1] \end{bmatrix} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix} x[0]$$

$$\mathbf{w} = \hat{Q} x[0]$$

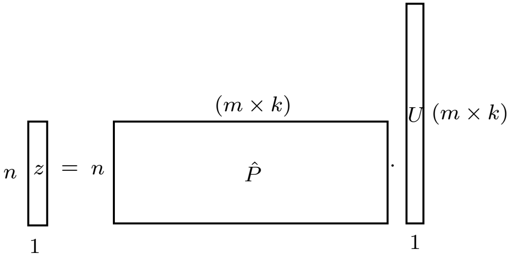
$\mathbf{w} \in \mathbb{R}^{km \times 1}, \hat{Q} \in \mathbb{R}^{km \times n}$. Usually, we have $km > n$.

To uniquely solve $x[0]$, we need to have $\text{rank}[\hat{Q}] = \text{rank}[\hat{Q} \vdots \mathbf{w}] = n$. C-H $\Rightarrow k \to n$

Summary: Controllability

$$x[k] - A^k x[0] = [B \vdots AB \vdots \cdots \vdots A^{k-1}B] \begin{bmatrix} u[k-1] \\ u[k-2] \\ \vdots \\ u[0] \end{bmatrix}$$

$$z = \hat{P}U$$



Overactuated system $m \times k > n$
Want to calculate U for arbitrary z (or, if u is not unique)

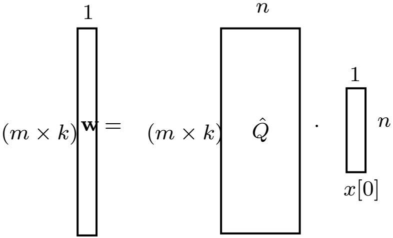
Let $W = [\hat{P} \vdots z]$ $\quad rank(\hat{P})$ must equal to $rank(W)$
$\Rightarrow \hat{P}$ need to have full rank, i.e. n independent columns.
$\Rightarrow rank(\hat{P}) = n$
C.H $\Rightarrow rank(\hat{P}) = rank(P)$, so we need $rank(P) = n$

Summary: Observability

$$w\left[k\right]=y\left(k\right)-\sum_{m=0}^{k-1}CA^{k-m-1}Bu\left(m\right)-Du\left(k\right)\in\mathbb{R}^{m\times 1}$$

$$\begin{bmatrix} w\left[0\right] \\ w\left[1\right] \\ \vdots \\ w\left[k-1\right] \end{bmatrix}=\begin{bmatrix} C \\ CA \\ \vdots \\ CA^{k-1} \end{bmatrix}x\left[0\right]$$

$$w=\hat{Q}\cdot x[0]$$



Underactuated system

Because $w[k]$ is calcualted from $x[0]$ via s-s equation. We should always have a solution. $\Rightarrow rank(\left[\hat{Q}\vdots\mathbf{w}\right])=rank(\hat{Q})$. We need to get a unique solution for $x[0]$. Therefore, require $rank(\hat{Q})=n$. C-H $\Rightarrow rank(Q)=n$

Controllability & Observability for Continuous Time LTI Systems

Good news! It has the same as formulae as for the DT.

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

- is controllable $\Leftrightarrow$ rank$(P = [B \vdots AB \vdots \cdots \vdots A^{n-1}B]) = n$
- is observable $\Leftrightarrow$ rank$\left(Q = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}\right) = n$

Proof

$x(t_1) = e^{A(t_1-t_0)}x(t_0) + \int_{t_0}^{t_1} e^{A(t_1-\tau)}Bu(\tau)\,d\tau$

$\Rightarrow \int_{t_0}^{t_1} e^{A(t_1-\tau)}Bu(\tau)\,d\tau \ = \ x(t_1) - e^{A(t_1-t_0)}x(t_0)$

C-H: $e^{A(t_1-\tau)} = \sum_{i=1}^{n} \alpha_i(\tau)\,A^{n-i}$

$\Rightarrow \int_{t_0}^{t_1} e^{A(t_1-\tau)}Bu(\tau)\,d\tau = \int_{t_0}^{t_1} \left[\sum_{i=1}^{n} \alpha_i(\tau)\,A^{n-i}B\right] u(\tau)\,d\tau$

$= \int_{t_0}^{t_1} \left[A^{n-1}B\alpha_1(\tau)\,u(\tau) + A^{n-2}B\alpha_2(\tau)\,u(\tau) + \cdots + B\alpha_n(\tau)\,u(\tau)\right] d\tau$

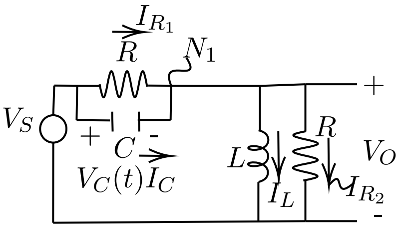
$$= A^{n-1}B \underbrace{\int_{t_0}^{t_1} \alpha_1(\tau)\,u(\tau)\,d\tau}_{\beta_1} + A^{n-2}B \underbrace{\int_{t_0}^{t_1} \alpha_2(\tau)\,u(\tau)\,d\tau}_{\beta_2} + \cdots + B \underbrace{\int_{t_0}^{t_1} \alpha_n(\tau)\,u(\tau)\,d\tau}_{\beta_n}$$

$$= \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} \begin{bmatrix} \beta_n \\ \beta_{n-1} \\ \vdots \\ \beta_1 \end{bmatrix} = x(t_1) - e^{A(t-t_0)}x(t_0)$$

To make $\beta_1 \cdots \beta_n$ solvable for $\forall x(t_0)$ and $x(t_1) \Rightarrow r(P) = n$

Example

For what values of R is the system $\dot{x} = Ax + Bu$ uncontrollable and unobservable?



Let $x_1 = V_C$ and $x_2 = I_L$

$u = V_S$, $y = V_O$

$$V_C = V_S - V_O = V_S - L\frac{dI_L}{dt} \Rightarrow \frac{dI_L}{dt} = \frac{1}{L}V_S - \frac{1}{L}V_C$$

$$\frac{dV_C}{dt} = \frac{1}{C}I_C = \frac{1}{C}(I_L + I_{R_2} - I_{R_1})$$

$$= \frac{1}{C}(I_L + \frac{V_S - V_C}{R} - \frac{V_C}{R}) = \frac{I_L}{C} - \frac{2V_C}{RC} + \frac{V_S}{RC}$$

$$\Rightarrow \dot{X} = \begin{bmatrix} -\frac{2}{RC} & \frac{1}{C} \\ -\frac{1}{L} & 0 \end{bmatrix} X + \begin{bmatrix} \frac{1}{RC} \\ \frac{1}{L} \end{bmatrix} u, X = [V_C, I_L]^T$$

$$Y = \begin{bmatrix} -1 & 0 \end{bmatrix} X + 1 \cdot u$$

Example cont.

$$P = \begin{bmatrix} B & AB \end{bmatrix} = \begin{bmatrix} \frac{1}{RC} & -\frac{2}{R^2C^2} + \frac{1}{LC} \\ \frac{1}{L} & -\frac{1}{RLC} \end{bmatrix}$$

To test rank, look at $|P| = 0$

$$|P| = \frac{1}{R^2LC^2} - \frac{1}{L^2C} = 0 \Rightarrow R = \sqrt{\frac{L}{C}} \text{ to lose controllability}$$

$$Q = \begin{bmatrix} C \\ CA \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ \frac{2}{RC} & -\frac{1}{C} \end{bmatrix}$$

$r(Q) = 2, \forall R \Rightarrow$ always observable

Controllability/Observability Remains under a Similarity Transforamation

$\dot{x} = Ax + Bu$. Let $x = M\hat{x}$
$\Rightarrow M\dot{\hat{x}} = AM\hat{x} + Bu \Rightarrow \dot{\hat{x}} = M^{-1}AM\hat{x} + M^{-1}Bu$
$\Rightarrow \hat{A} = M^{-1}AM$ and $\hat{B} = M^{-1}B$
Then

$$\begin{aligned}\tilde{P} &= \begin{bmatrix}\hat{B} & \hat{A}\hat{B} & \cdots & \hat{A}^{n-1}\hat{B}\end{bmatrix} \\ &= \begin{bmatrix}M^{-1}B & M^{-1}AMM^{-1}B & M^{-1}AMM^{-1}AMM^{-1}B & \cdots\end{bmatrix} \\ &= M^{-1}P\end{aligned}$$

Because for any A and B: $\mathrm{rank}(AB) \leq \min(\mathrm{rank}(A), \mathrm{rank}(B))$
$\mathrm{rank}(P) = \mathrm{rank}(M\tilde{P}) \leq \min(\mathrm{rank}(M), \mathrm{rank}(\tilde{P})) \Rightarrow n \leq \min(n, \mathrm{rank}(\tilde{P})) \Rightarrow \mathrm{rank}(\tilde{P}) = n$

Table of Contents

1 Controllability & Observability Matrices
- Solutions of Simultaneous Linear Equations

2 Popov-Belevitch-Hautus Tests

3 Jordan Form Tests
- Singular Value Decomposition
- Gram-Schmidt Process
- QR Decomposition

Popov-Belevitch-Hautus Tests (PBH)

Popov-Belevitch-Hautus Tests-Controllable

An LTI system is uncontrollable $\Leftrightarrow \exists$ left e-vector v, i.e. $v \neq 0$, $vA = \lambda v$ for e-value λ, s.t. $vB = 0$.

Note: Needed later on for Jordan form!

Popov-Belevitch-Hautus Tests (PBH)

An LTI system is uncontrollable $\Leftrightarrow \exists v \neq 0$, s.t. $vA = \lambda v$ for left λ and $vB = 0$.
A **left eigenvector** of A is a vector, s.t. $v \in \mathbb{C}^{1\times n}$, $vA = \lambda v$.

- Proof:

$S(\Leftarrow)$:Suppose exists such a v

Then $vAB = \lambda vB = 0, vA^2B = \lambda^2 vB = 0, \cdots, vA^{n-1}B = \lambda^{n-1}vB = 0$

$\Rightarrow v\begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} = 0 \Rightarrow r(P) < n \Rightarrow$ uncontrollable.

$N(\Rightarrow)$:Assume $r(P) < n$.

Then $\exists v \in \mathbb{R}^{1\times n} \neq 0$ s.t. $vP = 0$

$\Rightarrow \begin{bmatrix} vB & vAB & \cdots & vA^{n-1}B \end{bmatrix} = 0.v$ is a **left eigenvector**

$\Rightarrow \begin{bmatrix} vB & \lambda vB & \cdots & \lambda^{n-1}vB \end{bmatrix} = 0 \Rightarrow vB = 0.$

Popov-Belevitch-Hautus Tests (PBH)

Popov-Belevitch-Hautus Tests- Controllable

An LTI system is uncontrollable $\Leftrightarrow \exists$ left e-vector v, i.e. $v \neq 0$, $vA = \lambda v$ for e-value λ, s.t. $vB = 0$.

An LTI system is controllable $\Leftrightarrow \text{rank}([\lambda I - A \vdots B]) = n$, $\forall \lambda$ a eigenvalue of A.

Popov-Belevitch-Hautus Tests (PBH)

- Proof:

$$S(\Leftarrow) : r([\lambda I - A \vdots B]) = n$$

$$\Rightarrow \text{There does not exist } v \neq 0 \text{ s.t. } v[\lambda I - A \vdots B] = [v(\lambda I - A) \vdots vB] = 0$$

$$\Rightarrow \text{There does not exist } v \neq 0 \text{ s.t. } vA = \lambda A \text{ and } vB = 0$$

$$N(\Rightarrow) : \text{Follows reverse of above}$$

Popov-Belevitch-Hautus Tests (PBH)

The analogous statement is true for observability.

Popov-Belevitch-Hautus Tests Observability

An LTI system is unobservable $\Leftrightarrow \exists v \neq 0$ s.t. $Av = \lambda v$ and $Cv = 0$.

An LTI system is observable $\Leftrightarrow r\left(\begin{bmatrix} \lambda I - A \\ C \end{bmatrix}\right) = n$, $\forall \lambda$ an eigenvalue of A.

Recap: Controllability & Observability

- **Controllability**:
  A system is controllable if $\exists u(t), t \in [t_0, t_1]$ that transfers the system from any $x(t_0)$ to any $x(t_1)$.
- **Observability**:
  A system is observable if knowing $u(t), y(t), t \in [t_0, t_1]$ is sufficient to uniquely solve for $\forall x(t_0)$.

Recap: Test Controllability & Observability

Test controllability using $\text{rank}(P)$

A DT LTI system is controllable $\Leftrightarrow \text{rank}(P) = n$, where

$$P = [B \vdots AB \vdots \cdots \vdots A^{n-1}B]$$

Test observability using $\text{rank}(Q)$

A DT LTI system is observable $\Leftrightarrow \text{rank}(Q) = n$, where

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

Controllability and Observability are not changed under a similarity transformation.

Recap: Popov-Belevitch-Hautus Tests (PBH)

Popov-Belevitch-Hautus Tests: Controllable

An LTI system is uncontrollable $\Leftrightarrow \exists v \neq 0$, $vA = \lambda v$, s.t. $vB = 0$.

An LTI system is controllable $\Leftrightarrow \text{rank}([\lambda I - A \vdots B]) = n$, $\forall \lambda$ an eigenvalue of A.

Popov-Belevitch-Hautus Tests: Observability

An LTI system is unobservable $\Leftrightarrow \exists v \neq 0$ s.t. $Av = \lambda v$ and $Cv = 0$.

An LTI system is observable $\Leftrightarrow r\left(\begin{bmatrix} \lambda I - A \\ C \end{bmatrix}\right) = n$, $\forall \lambda$ an eigenvalue of A.

Table of Contents

1 Controllability & Observability Matrices
- Solutions of Simultaneous Linear Equations

2 Popov-Belevitch-Hautus Tests

3 Jordan Form Tests
- Singular Value Decomposition
- Gram-Schmidt Process
- QR Decomposition

Jordan Form

We can also use the Jordan form $J = M^{-1}AM$ and $\hat{B} = M^{-1}B$ to test the controllability/observability. Organize the Jordan form s.t. all Jordan blocks with the same eigenvalues are adjacent.

$$J = \left[\begin{array}{ccc|c|ccc} J_{11}^{\lambda_1} & & & & & & \\ & \ddots & & & & & \\ & & J_{1q_1}^{\lambda_1} & & & & \\ \hline & & & \ddots & & & \\ \hline & & & & J_{p1}^{\lambda_p} & & \\ & & & & & \ddots & \\ & & & & & & J_{pq_p}^{\lambda_p} \end{array}\right], \hat{B} = \left[\begin{array}{c} \hat{B}_{11}^{\lambda_1} \\ \vdots \\ \hat{B}_{1q_1}^{\lambda_1} \\ \hline \\ \hline \hat{B}_{p1}^{\lambda_p} \\ \vdots \\ \hat{B}_{pq_p}^{\lambda_p} \end{array}\right]$$

where q_i the number of Jordan blocks associated with λ_i and p the number of distinct eigenvalues. Now look at the PBH test, i.e. $r([\lambda I - J \vdots B])$. Note that for blocks with $\lambda_i \neq \lambda_j$, $[\lambda_i I - J_j]$ has full rank.

Example

Check blocks associated with λ_i , e.g.

$$J = \begin{bmatrix} \lambda_i & 0 & 0 \\ 0 & \lambda_i & 1 \\ 0 & 0 & \lambda_i \end{bmatrix} \xrightarrow{\lambda_i I - J} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow r\left(\lambda_i I - J\right) = 1$$

$$\Rightarrow \left[\lambda I - J, \hat{B}\right] = \begin{bmatrix} 0 & 0 & 0 & \widehat{B}_1 \\ 0 & 0 & -1 & \widehat{B}_2 \\ 0 & 0 & 0 & \widehat{B}_3 \end{bmatrix}$$

$$\Longrightarrow \quad \text{for} \quad r\left(\left[\lambda I - J, \hat{B}\right]\right) = 3 \Leftrightarrow r\left(\begin{bmatrix} \widehat{B}_1 \\ \widehat{B}_3 \end{bmatrix}\right) = 2$$

Intuitively, we can pass the influence via the "1" between states, if not, then B needs help influence different states independently.

Test Controllability using Jordan Form

Theorem

Let $\hat{B}^{\lambda_i}$ be the matrix of rows of $\hat{B}$ corresponding to the last row of each Jordan block corresponding to λ_i. Then an LTI system is controllable $\Leftrightarrow$ $\hat{B}^{\lambda_i}$ has full row rank for any λ_i.

- For a system with one input channel $\hat{B}$ is a column vector $\Rightarrow$each eigenvalue can only have 1 Jordan block to be controllable.
  $\dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u, J = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \Rightarrow$ Not controllable
- For A with distinct eigenvalues $\Rightarrow$ rows of $\hat{B}$ just need be non-zero.
  $\dot{x} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u, J = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \Rightarrow$ Not controllable

Test Observability using Jordan Form

Theorem

Let $\hat{C}^{\lambda_i}$ be the matrix of columns of $\hat{C}$ corresponding to the first columns of each Jordan block corresponding to λ_i. Then an LTI system is observable$\Leftrightarrow\hat{C}^{\lambda_i}$ has full column rank for any λ_i

- For a system with one observation channel $\hat{C}$ is a row vector $\Rightarrow$each eigenvalue can only have 1 Jordan block to be observable.
- For A with distinct eigenvalues $\Rightarrow$ columns of $\hat{C}$ just need be non-zero.

Example

$$\hat{A} = \begin{bmatrix} -5 & & & & & & & \\ & -5 & 1 & & & & & \\ & 0 & -5 & & & & & \\ & & & 3 & & & & \\ & & & & -4 & 1 & & \\ & & & & 0 & -4 & & \\ & & & & & & 0 & \\ & & & & & & & 0 \end{bmatrix}, \hat{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 2 & 2 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

$$\hat{C} = \begin{bmatrix} -1 & 1 & -1 & -2 & 1 & 0 & 0 & 3 \end{bmatrix}$$

Example

$$\hat{B}^{-5} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \Rightarrow r\left(\hat{B}^{-5}\right) = 2$$

$$\hat{B}^{3} = \begin{bmatrix} 2 & 2 \end{bmatrix}$$

$$\hat{B}^{-4} = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

$$\hat{B}^{0} = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \Rightarrow r\left(\hat{B}^{0}\right) = 1 \ (not\ controllable)$$

$$\hat{C}^{-5} = \begin{bmatrix} -1 & 1 \end{bmatrix} \Longrightarrow r\left(\hat{C}^{-5}\right) = 1 \ (not\ observable)$$

$$\hat{C}^{3} = \begin{bmatrix} -2 \end{bmatrix}$$

$$\hat{C}^{-4} = \begin{bmatrix} 1 \end{bmatrix}$$

$$\hat{C}^{0} = \begin{bmatrix} 0 & 3 \end{bmatrix} \ (not\ observable)$$

modes $\lambda = -5, -4, 3$ controllable; $\lambda = 0$ not; modes $\lambda = -4, 3$ observable; $\lambda = -5, 0$ not

How Do Computers Calculate Rank

- Gaussian Elimination is good to manually check the rank. But there is a huge risk to use computer to blindly do it due to the computational error.

  Example: The rank of $\mathbf{A} = \begin{bmatrix} 1 & 10^{10} \\ 0 & 1 \end{bmatrix} \Rightarrow$ rank=2

  Calculate rank of $\mathbf{A} + \mathbf{E}$ where $\mathbf{E}$ (computational error) $= \begin{bmatrix} 0 & 0 \\ 10^{-10} & 0 \end{bmatrix}$

  $r(\begin{bmatrix} 1 & 10^{10} \\ 10^{-10} & 1 \end{bmatrix}) = 1 \Rightarrow$ rank=1, which means $\mathbf{A}$ is very close to a defective matrix!
- We need a method that not only tells us whether a matrix is defective or not but also assess how close it is to be defective. $\Rightarrow$ **Singular Value Decomposition (SVD)**

Singular Value Decomposition

$\forall \mathbf{A} \in \mathbb{R}^{(m\times n)}$ we can do the **Singular Value Decomposition** s.t.

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T = \begin{bmatrix} \\ \mathbf{u}_1 & \cdots & \mathbf{u}_r & \cdots & \mathbf{u}_m \\ \\ \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & & & \\ & \ddots & & \mathbf{0} \\ & & \sigma_r & \\ \hline & \mathbf{0} & & \mathbf{0} \end{array}\right] \begin{bmatrix} \\ \mathbf{v}_1 & \cdots & \mathbf{v}_r & \cdots & \mathbf{v}_n \\ \\ \end{bmatrix}^T$$

$\mathbf{U} \in \mathbb{R}^{m\times m}, \boldsymbol{\Sigma} \in \mathbb{R}^{m\times n}, \mathbf{V} \in \mathbb{R}^{n\times n}$ A few nice features:

- It is canonical/fixed decomposition (compared to Jordan form). No ambiguity
- It can apply ANY matrix (similarity decomposition only applies to square matrix)
- σ_i are all positive real number: $\sigma_1 \geq \sigma_2 \geq \sigma_3 \cdots \geq \sigma_r \geq 0$. Good to measure a degree, e.g. defectiveness.
- $\mathbf{A}^T\mathbf{A}\mathbf{v}_i = \sigma_i^2\mathbf{v}_i$, $\mathbf{A}\mathbf{A}^T\mathbf{u}_i = \sigma_i^2\mathbf{u}_i$ $\mathbf{u}_i \Rightarrow \mathbf{v}_i$ are the e-vectors of $\mathbf{A}^T\mathbf{A}$, $\mathbf{u}_i$ are the e-vectors of $\mathbf{A}\mathbf{A}^T$. They share the same e-values σ_i^2
- $\mathbf{v}_i^T\mathbf{v}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}, \qquad \mathbf{u}_i^T\mathbf{u}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \Rightarrow \mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{I}$

How Does Singular Values Deal with Numerical Issues

The rank of $\mathbf{A} = \begin{bmatrix} 1 & 10^{10} \\ 0 & 1 \end{bmatrix} \Rightarrow$ rank=2

Calculate rank of $\mathbf{A} + \mathbf{E} = \begin{bmatrix} 1 & 10^{10} \\ 10^{-10} & 1 \end{bmatrix}$ where $\mathbf{E} = \begin{bmatrix} 0 & 0 \\ 10^{-10} & 0 \end{bmatrix} \Rightarrow$ rank=1

Check the singular values of $\mathbf{A}$

$\sigma_1(\mathbf{A}) = 10^{10}$, $\sigma_2(\mathbf{A}) \approx 10^{-10}$.

We can see that the smallest sigular value is already very close to 0, which robustly implies the matrix is close to be defective.

Symmetric Matrix

In order to learn SVD, we need to under understand **positive definite** and **positive definite**. Given $\mathbf{A}$ real, it is a symmetric matrix: $\mathbf{A}^T = \mathbf{A}$. Symmetric matrices have very nice features:

- All its eigenvalues are real

  $$\mathbf{v}^H\mathbf{A}\mathbf{v} = \mathbf{v}^H\lambda\mathbf{v} = \lambda\mathbf{v}^H\mathbf{v}, \begin{cases} (\mathbf{v}^H\mathbf{A}\mathbf{v})^H = \mathbf{v}^H\mathbf{A}^H\mathbf{v} = \mathbf{v}^H\mathbf{A}^T\mathbf{v} = \mathbf{v}^H\mathbf{A}\mathbf{v} \\ (\mathbf{v}^H\mathbf{v})^H = \mathbf{v}^H\mathbf{v} \end{cases}.$$

  Both $\mathbf{v}^H\mathbf{A}\mathbf{v}$ and $\lambda\mathbf{v}^H\mathbf{v}$ are real, λ is real.

Symmetric Matrix (cont.)

Given $\mathbf{A}$ real, it is a symmetric matrix: $\mathbf{A}^T = \mathbf{A}$.

- $\mathbf{A}$ has n orthogonal eigenvectors
  Let $\lambda_1 \neq \lambda_2, \mathbf{A}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1,$
  $\Rightarrow \mathbf{A}\mathbf{v}_2 = \lambda_2 \mathbf{v}_2$. Multiply $\mathbf{v}_2^T$ on both side $\Rightarrow \mathbf{v}_2^T \mathbf{A} \mathbf{v}_1 = \lambda_1 \mathbf{v}_2^T \mathbf{v}_1$.
  Take transpose $\mathbf{v}_1^T \mathbf{A}^T \mathbf{v}_2 = \lambda_2 \mathbf{v}_2^T \mathbf{v}_1$. Since $\mathbf{A}$ is symmetric: $\mathbf{v}_1^T \mathbf{A} \mathbf{v}_2 = \lambda_2 \mathbf{v}_2^T \mathbf{v}_1$
  $\Rightarrow \mathbf{v}_1^T \lambda_2 \mathbf{v}_2 = \lambda_1 \mathbf{v}_2^T \mathbf{v}_1, (\lambda_1 - \lambda_2)\, \mathbf{v}_1^T \mathbf{v}_2 = 0, \lambda_1 \neq \lambda_2 \Rightarrow \mathbf{v}_1^T \mathbf{v}_2 = 0$
- Let $\mathbf{M} = [\mathbf{v}_1, \cdots, \mathbf{v}_n]$, where $\mathbf{v}_i$ are normalized e-vectors (orthonormal e-vectors), i.e. $||\mathbf{v}_i||_2 = 1 \Rightarrow \mathbf{M}^T\mathbf{M} = \mathbf{I}$, $\mathbf{M}^T = \mathbf{M}^{-1} \Rightarrow \mathbf{A} = \mathbf{M}\Lambda\mathbf{M}^{-1} = \mathbf{M}\Lambda\mathbf{M}^T$,
  where Λ is diagonal with real values and columns of $\mathbf{M}$ are orthogonal.

Great! Now we just need to construct a symmetric matrix related $\mathbf{A}$ to help us to get these nice features. The easiest one is $\mathbf{A}^T\mathbf{A}$! Actually it is a special type of symmetric matrix: positive definite matrix!

Positive Definiteness

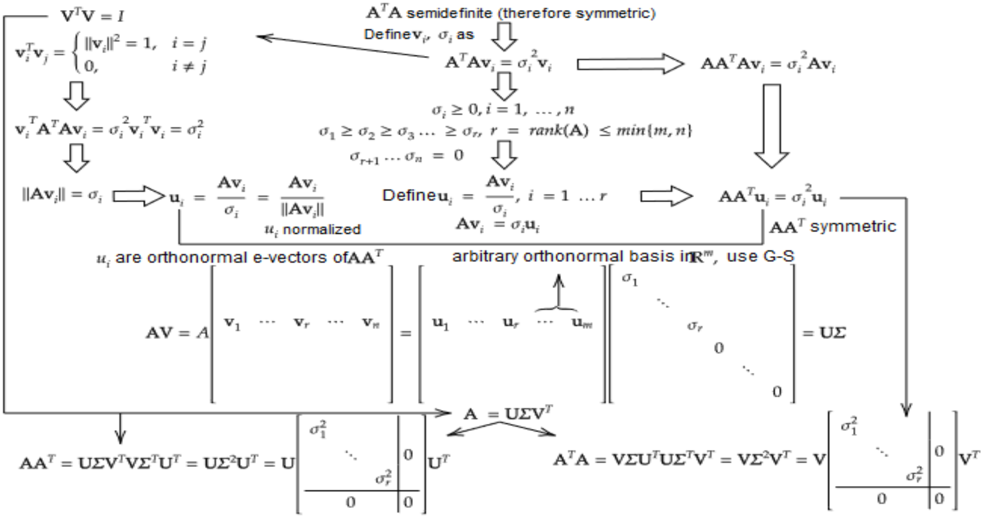
- A symmetric matrix $\mathbf{P}$ is **positive definite** if for $\forall \mathbf{x} \neq 0$, $\mathbf{x}^T\mathbf{P}\mathbf{x} > 0$.
  Every eigenvalue of $\mathbf{P}$ is positive
  $\mathbf{x}^T\mathbf{P}\mathbf{x} = \mathbf{x}^T\mathbf{M}\hat{\mathbf{P}}\mathbf{M}^T\mathbf{x} = (\mathbf{M}^T\mathbf{x})^T\hat{\mathbf{P}}(\mathbf{M}^T\mathbf{x}) = \mathbf{z}^T\hat{\mathbf{P}}\mathbf{z} = \sum \lambda_{ii} z_i^2 \geq 0$
  where $\mathbf{z} = \mathbf{M}^T\mathbf{x}$
- A symmetric matrix $\mathbf{P}$ is **positive semi-definite** if for $\forall \mathbf{x} \neq 0$, $\mathbf{x}^T\mathbf{P}\mathbf{x} \geq 0$.
  Every eigenvalue of $\mathbf{P}$ is nonnegative

$\mathbf{A}^T\mathbf{A}$ is positive semi-definite, because

- Symmetric: $(\mathbf{A}^T\mathbf{A})^T = \mathbf{A}^T\mathbf{A}$
- $\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x} = (\mathbf{A}\mathbf{x})^T(\mathbf{A}\mathbf{x}) \geq 0$

Similarly, $\mathbf{A}\mathbf{A}^T$ is positive semi-definite too.

Derivation



$V^TV = I$
$v_i^Tv_j = \begin{cases} \|v_i\|^2 = 1, & i = j \\ 0, & i \neq j \end{cases}$
$v_i^TA^TAv_i = \sigma_i^2 v_i^Tv_i = \sigma_i^2$
$\|Av_i\| = \sigma_i$
A^TA semidefinite (therefore symmetric)
Define v_i, σ_i as
$A^TAv_i = \sigma_i^2 v_i$
$AA^TAv_i = \sigma_i^2 Av_i$
$\sigma_i \geq 0, i = 1, \ldots, n$
$\sigma_1 \geq \sigma_2 \geq \sigma_3 \ldots \geq \sigma_r,\ r = rank(A) \leq min\{m, n\}$
$\sigma_{r+1} \ldots \sigma_n = 0$
$u_i = \frac{Av_i}{\sigma_i} = \frac{Av_i}{\|Av_i\|}$
u_i normalized
Define $u_i = \frac{Av_i}{\sigma_i},\ i = 1 \ldots r$
$Av_i = \sigma_i u_i$
$AA^Tu_i = \sigma_i^2 u_i$
AA^T symmetric
u_i are orthonormal e-vectors of AA^T
arbitrary orthonormal basis in $\mathbb{R}^m$, use G-S
$AV = A\begin{bmatrix} v_1 & \cdots & v_r & \cdots & v_n \end{bmatrix} = \begin{bmatrix} u_1 & \cdots & u_r & \cdots & u_m \end{bmatrix}\begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix} = U\Sigma$
$A = U\Sigma V^T$
$AA^T = U\Sigma V^TV\Sigma^TU^T = U\Sigma^2U^T = U\begin{bmatrix} \sigma_1^2 & & & 0 \\ & \ddots & & \\ & & \sigma_r^2 & \\ & 0 & & 0 \end{bmatrix}U^T$
$A^TA = V\Sigma U^TU\Sigma^TV^T = V\Sigma^2V^T = V\begin{bmatrix} \sigma_1^2 & & & 0 \\ & \ddots & & \\ & & \sigma_r^2 & \\ & 0 & & 0 \end{bmatrix}V^T$

Gram-Schmidt Process

Given a basis $\{\mathbf{y}_1, ..., \mathbf{y}_n\}$ find a new basis $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$ that is orthonormal and is a basis for span $\{\mathbf{y}_1, ..., \mathbf{y}_n\}$.

1. Generally: $\mathbf{v}_k = \mathbf{y}_k - \sum_{j=1}^{k-1} \frac{\langle \mathbf{v}_j, \mathbf{y}_k \rangle}{\|\mathbf{v}_j\|_2} \cdot \mathbf{v}_j$
2. Normalize $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ as $\{\frac{\mathbf{v}_1}{\|\mathbf{v}_2\|}, \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}, \ldots, \frac{\mathbf{v}_n}{\|\mathbf{v}_n\|}\}$ to obtain an orthonormal set.

Example

3×3 example:
Consider

$$\mathbf{A} = \left[\begin{array}{c|c|c} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{array}\right]$$

$$\mathbf{y}_1 \quad \mathbf{y}_2 \quad \mathbf{y}_3$$

Let $\mathbf{v}_1 = \mathbf{y}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. Applying Gram-schmidt process,

$$\mathbf{v}_2 = \mathbf{y}_2 - \frac{\langle \mathbf{v}_1, \mathbf{y}_2 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{v}_3 = \mathbf{y}_3 - \frac{\langle \mathbf{v}_1, \mathbf{y}_3 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\langle \mathbf{v}_2, \mathbf{y}_3 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Consider orthornormal basis $Q = \begin{bmatrix} \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} & \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} & \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} \end{bmatrix} = \left[\begin{array}{c|c|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right]$. Let $\mathbf{A} = \mathbf{QR}$

$$\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3$$

$$\mathbf{R} = \mathbf{Q}^{-1}\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

QR Decomposition

Orthornormal basis spanning $\mathbf{A} = \{\mathbf{y}_1, ..., \mathbf{y}_n\}$ can be computed using Gram-Schmidt process

$$\mathbf{v}_1 = \mathbf{y}_1 \qquad \mathbf{q}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}$$

$$\mathbf{v}_2 = \mathbf{y}_2 - \frac{\langle \mathbf{v}_1, \mathbf{y}_2 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 \qquad \mathbf{q}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|}$$

$$\mathbf{v}_3 = \mathbf{y}_3 - \frac{\langle \mathbf{v}_1, \mathbf{y}_3 \rangle}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 - \frac{\langle \mathbf{v}_2, \mathbf{y}_3 \rangle}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 \qquad \mathbf{q}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|}$$

$$\vdots \qquad \vdots$$

$$\mathbf{v}_k = \mathbf{y}_k - \sum_{j=1}^{k-1} \frac{\langle \mathbf{v}_j, \mathbf{y}_k \rangle}{\|\mathbf{v}_j\|^2} \qquad \mathbf{q}_k = \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|}$$

The $\mathbf{y}_i$s can be expressed over the newly computed orthonormal basis as

$$\mathbf{y}_1 = \langle \mathbf{q}_1, \mathbf{y}_1 \rangle \mathbf{q}_1$$

$$\mathbf{y}_2 = \langle \mathbf{q}_1, \mathbf{y}_2 \rangle \mathbf{q}_1 + \langle \mathbf{q}_2, \mathbf{y}_2 \rangle \mathbf{q}_2$$

$$\mathbf{y}_3 = \langle \mathbf{q}_1, \mathbf{y}_3 \rangle \mathbf{q}_1 + \langle \mathbf{q}_2, \mathbf{y}_3 \rangle \mathbf{q}_2 + \langle \mathbf{q}_3, \mathbf{y}_3 \rangle \mathbf{q}_3$$

$$\vdots$$

$$\mathbf{y}_k = \sum_{j=1}^{k} \langle \mathbf{q}_j, \mathbf{y}_k \rangle \mathbf{q}_j$$

This can be written in matrix form as $\mathbf{A} = \mathbf{QR}$ where

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1 & \cdots & \mathbf{q}_n \end{bmatrix}, \quad \mathbf{R} = \begin{pmatrix} \langle \mathbf{q}_1, \mathbf{y}_1 \rangle & \langle \mathbf{q}_1, \mathbf{y}_2 \rangle & \langle \mathbf{q}_1, \mathbf{y}_3 \rangle & \cdots \\ 0 & \langle \mathbf{q}_2, \mathbf{y}_2 \rangle & \langle \mathbf{q}_2, \mathbf{y}_3 \rangle & \cdots \\ 0 & 0 & \langle \mathbf{q}_3, \mathbf{y}_3 \rangle & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Recap: Test Controllability & Observability using Jordan Form

Theorem

Let $\hat{B}^{\lambda_i}$ be the matrix of rows of $\hat{B}$ corresponding to the last row of each Jordan block corresponding to λ_i. Then an LTI system is controllable $\Leftrightarrow$ $\hat{B}^{\lambda_i}$ has full row rank for any λ_i.

Theorem

Let $\hat{C}^{\lambda_i}$ be the matrix of columns of $\hat{C}$ corresponding to the first columns of each Jordan block corresponding to λ_i. Then a LTI system is observable$\Leftrightarrow$$\hat{C}^{\lambda_i}$ has full column rank for any λ_i