Controllability

Consider a state determined dynamical system with a transition map $s(t_1, t_0, x_0, u(\cdot))$ such that

$$x(t_1) = s(t_1, t_0, x(t_0), u(\cdot))$$

This can be continuous time or discrete time.

**Controllability:** *The dynamical system is said to be controllable on* $[t_0, t_1]$ *if* $\forall$ initial and final states $x_0, x_1$, $\exists u(\cdot)$ *so that* $s(t_1, t_0, x_0, u) = x_1$.

*It is said to be controllable at* $t_0$ *if* $\forall x_0, x_1$, $\exists t_1 \geq t_0$ *and* $u(\cdot) \in \mathcal{U}$ *so that* $s(t_1, t_0, x_0, u) = x_1$

A system that is not controllable probably means the followings:

- If state space system is realized from input-output data, then the realization is redundant (too many states)

- If states are meaningful (physical) variables that need to be controlled, then the design of the actuators are deficient
• The effect of control is limited. There is also a possibility of instability

Remarks:

• The followings are equivalent:

1. A system is controllable at \( t = t_0 \).
2. Any state can be transferred from any state at \( t = t_0 \) to any other state in finite time.
3. For each \( x_0 \in \Sigma \),

\[
s(\cdot, t_0, x_0, \cdot) : [t_0, \infty) \times u(\cdot) \mapsto s(t_1, t_0, x_0, u(\cdot))
\]

is surjective (onto).

• Controllability does not say that \( x(t) \) remains at \( x_1 \). e.g. for

\[
\dot{x} = Ax + Bu
\]

to make \( x(t) = x_1 \ \forall t \geq t_0 \), it is necessary that \( Ax_1 \in \text{Range}(B) \). This is generally not true.
Example - An Uncontrollable System

- An linear actuator pushing 2 masses on each end of the actuator in space.

\[ m_1 \ddot{x}_1 = u; \quad m_2 \ddot{x}_2 = -u \]

- State: \( X = [x_1, \dot{x}_1, x_2, \dot{x}_2]^T \).

- Action and reaction are equal and opposite, and act on different bodies.

- Momentum is conserved:

\[ m_1 \dot{x}_1 + m_2 \dot{x}_2 = \text{constant} = m_1 \dot{x}_1(t_0) + m_2 \dot{x}_2(t_0) \]

- E.g. if both masses are initially at rest: it is not possible to choose \( u(\cdot) \) to make \( \dot{x}_1 = 0, \dot{x}_2 = 1 \).
Preliminary properties

- Memoryless feedback:

\[ u(t) = v(t) - g(t, x(t)) \]

with \( v(t) \) being the new input.

- A system is controllable if and only if the system under memoryless feedback with \( v(t) \) as the input is controllable.
• **Dynamic state feedback:**

\[
\dot{z}(t) = \alpha(t, x(t), z(t))
\]

\[
u(t) = v(t) - g(t, x(t), z(t))
\]

where \(v(t)\) is the new input and \(z(t)\) is the state of the controller

- If a system under **dynamic feedback** is controllable then the original system is.
- The converse is not true, i.e. the original system is controllable does not necessarily imply that the system under dynamic state feedback is. Why? (What is the state is of the feedback system?)
Some notations

Let $L : \mathcal{X}(=\mathbb{R}^p) \rightarrow \mathcal{Y}(=\mathbb{R}^q)$ be a linear map: i.e.

$$L(x_1 + x_2) = Lx_1 + Lx_2$$

$\mathcal{R}(L) \subset \mathcal{Y}$ denotes the range of $L$, given by:

$$\mathcal{R}(L) = \{y \in \mathcal{Y} | y = Lx \text{ for some } x \in \mathcal{X}\}$$

$\mathcal{N}(L) \subset \mathcal{X}$ denotes the null space of $L$, given by:

$$\mathcal{N}(L) = \{x \in \mathcal{X} | Lx = 0\}$$

Note: Both $\mathcal{R}(L)$ and $\mathcal{N}(L)$ are linear subspaces, i.e.

\[ y_1, y_2 \in \mathcal{R}(L) \Rightarrow (y_1 + y_2) \in \mathcal{R}(L) \]

\[ x_1, x_2 \in \mathcal{N}(L) \Rightarrow (x_1 + x_2) \in \mathcal{R}(L) \]

$L$ is called onto or surjective if $\mathcal{R}(L) = \mathcal{Y}$ (everything is within range).

$L$ is called into or one-to-one (cf. many-to-one) if $\mathcal{N}(L) = \{\}$. Thus,

$$Lx_1 = Lx_2 \iff x_1 = x_2$$
Linear Systems

Continuous time system:

\[ \dot{x}(t) = A(t)x(t) + B(t)u(t) \quad x \in \mathbb{R}^n \]

\[ x(t_1) = \Phi(t_1, t_0)x_0 + \int_{t_0}^{t_1} \Phi(t_1, t)B(t)u(t)dt. \]

The reachability map on \([t_0, t_1]\) is defined to be:

\[ L_{r,[t_0,t_1]}(u(\cdot)) = \int_{t_0}^{t_1} \Phi(t_1, t)B(t)u(t)dt \]

Thus, it is controllable on \([t_0, t_1]\) if and only if \(L_{r,[t_0,t_1]}(u(\cdot))\) is surjective (onto).

Notice that \(L_{r,[t_0,t_1]}\) determines the set of states that can be reached from the origin at \(t = t_1\).

The study of the range space of the linear map:

\[ L_{r,[t_0,t_1]} : \{u(\cdot)\} \rightarrow \mathbb{R}^n \]

is central to the study of controllability.
**Proposition** If a continuous time, possibly time varying, linear system is controllable on $[t_0, t_1]$ then it is controllable on $[t_0, t_2]$ where $t_2 \geq t_1$.

**Proof:** To transfer from $x(t_0) = x_0$ to $x(t_2) = x_2$, define

$$x_1 = \Phi(t_1, t_2)x_2$$

Suppose that $u^*[t_0, t_1]$ transfers $x(t_0) = x_0$ to $x(t_1) = x_1$. Choose

$$u(t) = \begin{cases} u^*(t) & t \in [t_0, t_1] \\ 0 & t \in (t_1, t_2) \end{cases}$$

This result does not hold for $t_0 < t_2 < t_1$. e.g. The system

$$\dot{x}(t) = B(t)u(t) \quad B(t) = \begin{cases} 0 & t \in [t_0, t_2] \\ I_n & t \in (t_2, t_1) \end{cases}$$

is controllable on $[t_0, t_1]$ but $x(t_2) = x(t_0)$ for any control $u(\cdot)$. 
Reduction Theorem

The followings are equivalent for a linear time varying differential system:

\[
\dot{x} = A(t)x + B(t)u
\]

1. It is controllable on \([t_0, t_1]\).

2. It is controllable to 0 on \([t_0, t_1]\), i.e. \(\forall x(t_0) = x_0\), there exists \(u(\tau), \tau \in [t_0, t_1]\) such that final state is \(x(t_1) = 0\).

3. It is reachable from \(x(t_0) = 0\), i.e. \(\forall x(t_1) = x_f\), there exists \(u(\tau), \tau \in [t_0, t_1]\) such that for \(x(t_0) = 0\), the final state is \(x(t_1) = x_f\).

Proof:

- Clearly, (1) \(\Rightarrow\) (2) and (3).

- To prove (2) \(\Rightarrow\) (1) and (3) \(\Rightarrow\) (1), use

\[
x(t_1) = \Phi(t_1, t_0)x(t_0) + \int_{t_0}^{t_1} \Phi(t_1, \tau)B(\tau)u(\tau)d\tau.
\]
and \( \Phi(t_1, t_0) \) is invertible, to construct the appropriate \( x(t_0) \) and \( x(t_f) \).
Controllability map: \( L_{c,[t_0,t_1]} \)

The controllability (to zero) map on \([t_0, t_1]\) is the map between \( u(\cdot) \) to the initial state \( x_0 \) such that \( x(t_1) = 0 \).

\[
L_{c,[t_0,t_1]}(u(\cdot)) = -\int_{t_0}^{t_1} \Phi(t_0, t)B(t)u(t)dt
\]

Proof:

\[
0 = x(t_1) = \Phi(t_1, t_0)x_0 + \int_{t_0}^{t_1} \Phi(t_1, \tau)B(\tau)u(\tau)d\tau
\]

\[
x_0 = -\Phi(t_1, t_0)^{-1}\int_{t_0}^{t_1} \Phi(t_1, \tau)B(\tau)u(\tau)d\tau
\]

\[
x_0 = -\int_{t_0}^{t_1} \Phi(t_0, t_1)\Phi(t_1, \tau)B(\tau)u(\tau)d\tau
\]

\[
x_0 = -\int_{t_0}^{t_1} \Phi(t_0, \tau)B(\tau)u(\tau)d\tau
\]
Discrete time system

\[ x(k + 1) = A(k)x(k) + B(k)u(k), \quad x \in \mathbb{R}^n, u \in \mathbb{R}^m \]

\[ x(k_1) = \Phi(k_1, k_0)x(k_0) + \sum_{k=k_0}^{k_1-1} \Phi(k_1, k + 1)B(k)u(k) \]

Write:

\[ \sum_{k=k_0}^{k_1-1} \Phi(k_1, k + 1)B(k)u(k) = C(k_0, k_1)U \]

where \( C \in \mathbb{R}^{n \times (k_1-k_0)m} \), \( U = \begin{pmatrix} u(k_0) \\ \vdots \\ u(k_1 - 1) \end{pmatrix} \).

Thus, the system is controllable if and only if \( C(k_0, k_1) \) has rank \( n \).
Proposition Suppose that $A(k)$ is nonsingular for each $k_1 \leq k < k_2$, then the discrete time linear system is controllable on $[k_0, k_1]$ implies that it is controllable on $[k_0, k_2]$ where $k_2 \geq k_1$.

Proof: The proof is similar to the continuous case. However for a discrete time system, we need $A(k)$ nonsingular for $k_1 \leq k < k_2$ to ensure that $\Phi(k_2, k_1)$ is invertible. Q.E.D.
Example

Consider a unit point mass under control of a force:

\[
\frac{d}{dt} \begin{pmatrix} x \\ \dot{x} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ \dot{x} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u
\]

This is the same as \( \ddot{x} = u \).

Suppose that \( x(0) = \dot{x}(0) = 0 \), we would like to translate the mass to \( x(T) = 1, \dot{x}(T) = 0 \).

The reachability map \( L : u(\cdot) \rightarrow x(T) \) is

\[
x(T) = \int_0^T \Phi(T, t) B(t) u(t) dt
\]

Let use try to solve this problem using piecewise constant control:

\[
u(t) = \begin{cases} 
u_0 & 0 \leq t < T/10 \\ 
u_1 & T/10 \leq t < 2T/10 \\ \vdots \\ 
u_{10} & 9T/10 \leq t \leq T 
\end{cases}
\]

and find \( U = [u_1, u_2, \ldots, u_{10}]^T \).
The reachability map becomes:

\[ x(T) = (L_1 \ L_2 \ \ldots \ \ L_{10}) \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{10} \end{pmatrix} \]

where

\[ L_1 = \left[ \int_0^{T/10} \Phi(T, t) B(t) dt \right] = \frac{T}{10} \begin{pmatrix} T \\ 1 \end{pmatrix}, \]

\[ L_2 = \left[ \int_{T/10}^{2T/10} \Phi(T, t) B(t) dt \right] = \frac{T}{10} \begin{pmatrix} 0.9T \\ 1 \end{pmatrix}, \]

and \[ L_{10} = \frac{T}{10} \begin{pmatrix} 0.1T \\ 1 \end{pmatrix} \] etc.

Whether the matrix denoted by \( L_{pc} \in \mathbb{R}^{2 \times 10} \) is surjective (onto) (i.e. \( \mathcal{R}(L_{pc}) = \mathbb{R}^2 \)) determines whether we can achieve our task using piecewise constant control.

How do we check the column rank (number of independent columns) of a short and thin matrix??
Matrix rank

Proposition: Let $L \in \mathbb{R}^{n \times m}$,

$$\mathcal{R}(L) = \mathcal{R}(LL^T)$$

$$\mathcal{N}(L) = \mathcal{N}(L^TL)$$

Note:

- If $L \in \mathbb{R}^{n \times m}$ and $m > n$ (short and fat), then $LL^T \in \mathbb{R}^{n \times n}$ is both short and thin.

- To check whether the system is controllable we need to check if $LL^T$ is full rank - e.g. check its determinant.

- If $m < n$ (tall and thin), then $L^TL \in \mathbb{R}^{n \times n}$ is also both short and thin.

- Checking $\mathcal{N}(L^TL)$ will be useful when we talk about observability.
In our case, let $T = 10$ for convenience,

$$L_{pc} = \begin{pmatrix} 10 & 9 & \ldots & 1 \\ 1 & 1 & \ldots & 1 \end{pmatrix}$$

and

$$L_{pc}L_{pc}^T = \begin{pmatrix} 385 & 55 \\ 55 & 10 \end{pmatrix}$$

which is non-singular (full rank). Thus, the particle transfer problem is solvable using piecewise constant control.

If we make the number of pieces in the piecewise constant control larger and larger, then the multiplication of $L_{pc}$ by $L_{pc}^T$ becomes an integral:

$$L_{pc}L_{pc}^T = \int_0^T \Phi(T, t)B(t)B(t)^T\Phi(T, t)^T dt.$$
Reachability Grammian

For the system

\[ \dot{x} = A(t)x + B(t)u \]

The reachability grammian on the time interval \([t_0, t_1]\) is defined to be

\[
W_r = L_r L_r^* = \int_{t_0}^{t_1} \Phi(t_1, t) B(t) B(t)^T \Phi(t_1, t)^T dt
\]

Thus, controllability of a system on \([t_0, t_1]\) is equivalent to reachability grammian on the same time interval \([t_0, t_1]\) being full rank.
Controllability Grammian

For the system

\[ \dot{x} = A(t)x + B(t)u \]

The controllability grammian on the time interval \([t_0, t_1]\) is defined to be:

\[ W_c = L_c L_c^* = \int_{t_0}^{t_1} \Phi(t_0, t)B(t)B(t)^T \Phi(t_0, t)^T dt \]

Since controllability of a system on \([t_0, t_1]\) is equivalent to controllability to 0 which is determined by the controllability map being surjective, it is also equivalent to controllability grammian on the same time interval \([t_0, t_1]\) being full rank.
Least norm control - setup

- If reachability map $L_{r,[t_0,t_1]}$ is surjective, then there is at least one control input that can steer the system to the desired state.

- Typically, there are many such solutions. How does one find an appropriate one?

- One idea: minimize the cost of the control.

For illustration, consider the example of transferring a mass using piecewise constant inputs from initial state $x_0 = [x_s, 0]$ to a state $x_f = [x_d, 0]^T$. $L_r$ is given by the matrix,

$$L_{pc} = \begin{pmatrix} 10 & 9 & \ldots & 1 \\ 1 & 1 & \ldots & 1 \end{pmatrix}$$

and control input is:

$$U = [u(0), u(1), \ldots, u(9)]^T \in \mathbb{R}^{10}.$$
The control problem is to minimize the cost,

\[ J = \frac{1}{2} \sum_{k=0}^{9} u(k) R(k) u(k) = \frac{1}{2} U^T R_U U \]

where \( R(k) > 0 \) so that large \( R(k) \) implies control is expensive at time \( k \); subject to

\[ x_f - \Phi(10, 0) x_0 = L_{pc} U \]

\( R_U = \text{diag}(R(0), \ldots, R(9)) \) is a positive definite matrix.
Least Norm Control - solution

We use Lagrange multiplier method to perform the constrained minimization.

Let the augmented cost function be:

\[ J' = J(U) + \lambda^T (L_{pc}U - (x_f - \Phi(t_1, t_0)x_0)) \]

where \( \lambda \in \mathbb{R}^n \) is the Lagrange multiplier. The necessary condition for the constrained optimal solution \( U^* \) is for each component \( U_i \) of \( U \), and for each component \( \lambda_j \) of \( \lambda \),

\[ \frac{\partial J'}{\partial U_i} \bigg|_{U^*, \lambda^*} = 0 \quad \frac{\partial J'}{\partial \lambda_j} \bigg|_{U^*, \lambda^*} = 0 \]

These reduce to:

\[ R_U U^* + L_{pc}^T \lambda^* = 0 \]

\[ L_{pc}U^* = x_f - \Phi(t_1, t_0)x_0 \]

This gives

\[ x_f - \Phi(t_1, t_0)x_0 = L_{pc}U^* = L_{pc} \left( -R_U^{-1}L_{pc}^T \lambda^* \right) \]
so that

\[
\lambda^* = - \left( L_{pc} R_U^{-1} L_{pc}^T \right)^{-1} (x_f - \Phi(t_1, t_0)x_0)
\]

\[
U^* = R_U^{-1} L_{pc}^T \left( L_{pc} R_U^{-1} L_{pc}^T \right)^{-1} (x_f - \Phi(t_1, t_0)x_0).
\]
Geometric Interpretation

Let $x_0 = 0$ for simplicity. The least norm transfer problem is:

$$\min_U (J(U) = ) \langle U, U \rangle \quad \text{s.t.} \quad L_r U = x_f$$

where $\langle \cdot, \cdot \rangle$ is the inner product. The inner product can be defined in many ways, as long as:

1. $\langle u, v \rangle \in \mathbb{C}$ the complex field;
2. $\langle u, u \rangle \geq 0$.
3. For any $\alpha, \beta \in \mathbb{R}$,

   $$\langle u, \alpha v_1 + \beta v_2 \rangle = \alpha \langle u, v_1 \rangle + \beta \langle u, v_2 \rangle$$

4. $\langle u, v \rangle = \overline{\langle v, u \rangle}$

Some examples are:

1. for $u, v \in \mathbb{R}^n$, $\langle u, v \rangle = u^T v$;
2. \( <u, v> = u^T R v \) where \( R \in \mathbb{R}^{n \times n} \) is a positive definite matrix.

We say that \( U \) and \( V \) are normal to each other if \( <U, V> = 0 \).

Any solution that satisfies the constraint must be of the form

\[
(U - U^p) \in Null(L_r)
\]

where \( L_r U^p = x_f \) is any particular solution.

**Claim:** Let \( U^* \) be the optimal solution, and \( U \) is any solution that satisfies the constraint. Then, \( (U - U^*) \perp U^* \), i.e.

\[
< (U - U^*), U^* > = 0
\]

which is the normal equation for the least norm control problem.

**Proof:** Direct substitution.

\[
(U - U^*)^T R_U R_U^{-1} L_{pc}^T (L_{pc} R_U^{-1} L_{pc}^T)^{-1} x_f
\]

\[
= (L_{pc}(U - U^*))^T (L_{pc} R_U^{-1} L_{pc}^T)^{-1} x_f
\]

\[
= 0.
\]
Continuous Time Least Norm Control

\[ (L_{r,[t_0,t_1]}q)(t) = B^T(t)\Phi^T(t_1,t)q \]

\[ W_{r,[t_0,t_1]} = \left[ \int_{t_0}^{t_1} \Phi(t_1,\tau)B(\tau)B^T(\tau)\Phi(t_1,\tau)^T d\tau \right] \]

\[ u_{opt}(t) = B(t)^T\Phi(t_1,t)^TW_{r,[t_0,t_1]}^{-1}(x_1 - \Phi(t_1,t_0)x_0) \]

Eq. (1) solves \( u_{opt}(\cdot) \) in a batch form given \( x_0 \), i.e. it is open loop.
Recursive Least Norm Control

• Open loop control does not make use of feedback. Suppose now that \( x(t) \) is measured. Conceptually, we can solve (1) for \( u_{opt}(t) \) and let \( t_1 = t \).

• Computing \( W_{r,[t,t_1]} \) is expensive. Try recursion:

\[
\begin{align*}
    u_{opt}(t) &= B(t)^T \Phi(t_1, t)^T W_{r,[t,t_1]}^{-1} (x_1 - \Phi(t_1, t)x(t)) \\
    &= B(t)^T \Phi(t_1, t)^T W_{r,[t,t_1]}^{-1} x_1 \\
    &\quad - B(t)^T \Phi(t_1, t)^T W_{r,[t,t_1]}^{-1} \Phi(t_1, t)x(t) \\
    &= \alpha(t)x_1 - \beta(t)x(t)
\end{align*}
\]

where \( \alpha(t) \) and \( \beta(t) \) can be pre-computed, or computed recursively...

Notice that \( \alpha(t) \) and \( \beta(t) \) may not be well defined as \( t \to t_1 \).
For any invertible matrix \( Q(t) \in \mathbb{R}^{n \times n} \),
\[ Q(t)Q^{-1}(t) = I, \]
therefore
\[
\frac{d}{dt} Q^{-1}(t) = -Q^{-1}(t) \dot{Q}(t) Q^{-1}(t)
\]

Since
\[
W_{r,[t,t_1]} = \int_t^{t_1} \Phi(t_1, \tau) B(\tau) B(\tau)^* \Phi(t_1, \tau)^* \, d\tau
\]
\[
\frac{d}{dt} W_{r,[t,t_1]} = -\Phi(t_1, t) B(t) B(t)^* \Phi(t_1, t)^*
\]

- Thus we can solve for \( W_{r,[t,t_1]}^{-1} \) dynamically by integrating on-line
\[
\frac{d}{dt} W_{r,[t,t_1]}^{-1} = -W_{r,[t,t_1]}^{-1} \Phi(t_1, t) B(t) B(t)^* \Phi(t_1, t)^* W_{r,[t,t_1]}^{-1}
\]

- **Issue:** Does not work well when \( t \to t_1 \) because \( W_{r,[t,t_1]} \) can be close to singular.
Optimal Cost for Control

The **optimal cost** of the control for transferring a state from 0 to $x_d$ in the interval $[t_0, t_f]$ (assuming $t_f$ is fixed) is:

$$E_{\text{min}}(t_0, x_d) = J(U^*) = x_d^T W_r^{-1}[t_0, t_f] L_r L_r^T W_r^{-1}[t_0, t_f] x_d$$

$$= x_d^T W_r^{-1}[t_0, t_f] x_d$$

Thus, the inverse of the Reachability Grammian tells us how difficult it is to perform a state transfer from $x = 0$ to $x_d$. In particular, if $W_r$ is not invertible, for some $x_d$, the cost is infinite.
Cost as a function of initial time

Since

\[ W_{r,[t_0,t_f]} = \int_{t_0}^{t_f} \Phi(t_f, \tau) B(\tau) B(\tau)^* \Phi(t_f, \tau)^* d\tau \]

and the integrand is a positive semi-definite term.

Claim: For each \( x_d \), \( E_{min}(t_0, x_d) \) does not increase (typically decreases) as \( t_0 \) decreases (i.e. time interval increases):

\[ E_{min}(t_0, x_d) = x_d^T W_{r,[t_0,t_f]}^{-1} x_d \]

and consider

\[ \frac{d}{dt_0} E_{min}(t_0, x_d) = x_d^T \frac{d}{dt_0} W_{r,[t_0,t_f]}^{-1} x_d \]

\[ = W_{r,[t_0,t_f]}^{-1} \Phi(t_f, t_0) B(0) B(0)^* \Phi(t_f, t_0)^* W_{r,[t_0,t_f]}^{-1} \]
Thus,

\[
\frac{d}{dt_0} E_{\text{min}}(t_0, x_d) = x_d^T \left[ \frac{d}{dt_f} W_{r,[t_0,t_f]}^{-1} \right] x_d
\]

\[
= \left\| B(t_0)^* \Phi(t_f,t_0)^* W_{r,[t_0,t_f]}^{-1} x_d \right\|^2 \geq 0.
\]

Hence, the optimal control increases as the initial time increases (decreasing time interval).

In other words, the set of states reachable with a cost less than a give value grows as time interval increases.

If the system is unstable, in that, \( \Phi(t_f,t_0) \) becomes unbounded as \( t_0 \to -\infty \), then one can see that \( W_{r,[t_0,t_f]} \) also becomes unbounded. This means that for some \( x_d \), \( E_{\text{min}}(t_0, x_d) \to 0 \) and \( t_0 \to -\infty \). This means that some directions do not require any cost of control.

Which directions do these correspond to?
Cost as a function of final state

Since $W_{r,[t_0,t_f]}$ is symmetric and positive semi-definite, has $n$ orthogonal eigenvectors, $V = [v_1, \ldots, v_n]$, and associated eigenvalues $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$, such that $VV^T = I$. Then

$$W_{r,[t_0,t_f]} = V\Lambda V^T \Rightarrow W_{r,[t_0,t_f]}^{-1} = V \text{diag}(\lambda_1^{-1}, \ldots, \lambda_n^{-1})V^T$$

This shows that the most expensive direction to transfer to is the eigenvector associated with the minimum $\lambda_i$, and the least expensive direction is the eigenvector associated with the maximum $\lambda_i$. 
Linear Time Invariant Systems

Continuous time system:

\[ \dot{x} = Ax + Bu \]
\[ y = Cx \]  \hspace{1cm} (2)

Discrete time system:

\[ x(k + 1) = Ax(k) + Bu(k) \]
\[ y(k) = Cx(k) \]  \hspace{1cm} (3)

where \( x \in \mathbb{R}^n, u \in \mathbb{R}^m \).

**Note:** Because of time invariance, the system is controllable (observable) at \( t_0 \) means that it is controllable (observable) at any time.

We will develop tests for controllability and observability by using directly the matrices \( A \) and \( B \), \( A \) and \( C \).
Cost of control

The cost of transferring from $x = 0$ to $x = x_f$ at time $t_f$ is given by:

$$E_{\text{min}}(x_f, t_f) := \int_0^{t_f} u_{\text{opt}}^T(\tau)u_{\text{opt}}(\tau) d\tau = x_f^T W_r(t_f) x_f$$

**Example:** Consider a stable and an unstable system:

\begin{align*}
\ddot{x} + x + 2\dot{x} &= u \quad (4) \\
\ddot{x} + x - \dot{x} &= u \quad (5)
\end{align*}

Steer from $(x, \dot{x}) = (0, 0)$ to $(x, \dot{x}) = (1, 0)$.

![Figure 1: Min cost transfer for Stable system (Left) and Unstable system (Right)](image-url)
Controllability Tests

The three tests for controllability of system (2) is given by the following theorem.

**Theorem**  For the continuous time system (2), the followings are equivalent:

1. The system is controllable over the interval \([0, T]\), for some \(T > 0\).

2. The controllability gramian,

\[
W_{r,T} := \int_0^T \Phi(T,t)B(t)B(t)^T\Phi(T,t)^T dt
\]

\[
= \int_0^T e^{At}BB^*e^{A^*t} dt
\]

is full rank \(n\). (This test works for time varying systems also)

3. The controllability matrix

\[
C := \begin{pmatrix}
B & AB & A^2B & \cdots & A^{n-1}B
\end{pmatrix}
\]
has rank $n$. Notice that the controllability matrix has dimension $n \times nm$ where $m$ is the dimension of the control vector $u$.

4. (Popov-Belovich-Hautus (PBH) test) For each $s \in \mathbb{C}$, the matrix,

\[
\begin{pmatrix}
    sI - A & B
\end{pmatrix}
\]

has rank $n$.

Note that rank of $(sI - A, B)$ is less than $n$ only if $s$ is an eigenvalue of $A$. 

We need the Cayley-Hamilton Theorem to prove this result:

**Theorem** (Cayley Hamilton) Let \( A \in \mathbb{R}^{n \times n} \). Consider the polynomial

\[
\psi(s) = \det(sI - A)
\]

Then \( \psi(A) = 0 \).

**Proof**: This is true for any \( A \), but is easy to see using the eigen decomposition of \( A \) when \( A \) is semi-simple.

\[
A = T\Lambda T^{-1}
\]

where \( \Lambda \) is diagonal with eigenvalues on its diagonal. Since \( \psi(\lambda_i) = 0 \) by definition,

\[
\psi(A) = T\psi(\Lambda)T^{-1} = 0.
\]
Proof: Controllability test

(1) \iff (2): We have already seen before that controllability over the interval \([0, T]\) means that the range space of the reachability map,

\[
L_r(u) = \int_0^T e^{A(T-\tau)} Bu(\tau) d\tau
\]

must be the whole state space, \(\mathbb{R}^n\). Notice that

\[
W_{r,T} := \int_0^T e^{At} BB^* e^{A^*t} dt = L_r L_r^*
\]

where \(L_r^*\) is the adjoint (think transpose) of \(L_r\).

From the finite rank linear map theorem, \(\mathcal{R}(L_r) = \mathcal{R}(L_r L_r^*)\) but, \(L_r L_r^*\) is nothing but \(W_{r,T}\).

(2) \implies (3): We will show this by showing “not (3) \implies not (2)”. Suppose that (3) is not true. Then, there exists a \(1 \times n\) vector \(v^T\) so that,

\[
v^T B = v^T AB = v^T A^2 B \cdots = v^T A^{n-1} B = 0.
\]

Consider now,

\[
v^T W_{r,T} v = \int_0^T \|v^T e^{At} B\|_2^2 dt.
\]
Since
\[ e^{At} = I + At + \frac{A^2t^2}{2} + \frac{A^{n-1}t^{n-1}}{n-1!} + \cdots \]
and by the Cayley Hamilton Theorem, \( A^k \) is a linear combination of \( I, A, \ldots, A^{n-1} \), therefore,
\[ v^T e^{At} B = 0 \]
for all \( t \). Hence, \( v^T W_{r,T} v = 0 \) or \( W_{r,T} \) is not full rank.

(3) \( \Rightarrow \) (2): We will show this by showing “not (2) \( \Rightarrow \) not (3)”. If (2) is not true, then there exists \( 1 \times n \) vector \( v^T \) so that
\[ v^T W_{r,T} v = \int_0^T \| v^T e^{At} B \|_2^2 dt = 0. \]
Because \( e^{At} \) is continuous in \( t \), this implies that \( v^T e^{At} B = 0 \) for all \( t \in [0, T] \).

Hence, the all time derivatives of \( v^T e^{At} B = 0 \) for all
\[ t \in [0, T]. \text{ In particular, at } t = 0, \]
\[
\begin{align*}
v^T e^{At} B \bigg|_{t=0} &= v^T B = 0 \\
v^T \frac{d}{dt} e^{At} B \bigg|_{t=0} &= v^T AB = 0 \\
v^T \frac{d^2}{dt^2} e^{At} B \bigg|_{t=0} &= v^T A^2 B = 0 \\
&\vdots \\
v^T \frac{d^{n-1}}{dt^{n-1}} e^{At} B \bigg|_{t=0} &= v^T A^{n-1} B = 0
\end{align*}
\]

Hence, \[ v^T \begin{pmatrix} B & AB & A^2 B & \cdots & A^{n-1} B \end{pmatrix} = \begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix} \]
i.e. the controllability matrix of \((A, B)\) is not full rank.

(3) \(\Rightarrow\) (4): We will show “not (4) implies not (3)”. Suppose (4) is not true so that there exists a \(1 \times n\) vector \(v\), and \(\lambda \in \mathbb{C}\),
\[ v^T (\lambda I - A) = 0, \quad v^T B = 0. \]
for some $\lambda$. Then $v^T A = \lambda v^T$, $v^T A^2 = \lambda^2 v^T$ etc. Because $v^T B = 0$ by assumption, we have $v^T AB = \lambda v^T B = 0$, $v^T A^2 B = \lambda^2 v^T B = 0$ etc. Hence,

$$v^T B = v^T AB = v^T A^2 B = \cdots v^T A^{n-1} B = 0.$$ 

Therefore the controllability is not full rank.
Proof of (4) ⇒ (3) - Optional

To prove (4) ⇒ (3), we need the so called 2nd Representation Theorem. It is included here for completeness. We will not go through this part.

Definition Let $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $W \subset \mathbb{R}^n$ a subspace with the property that for each $w \in W, Aw \in W$. We say that $W$ is $A$–invariant.

Theorem (2nd Representation theorem) Let

- $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a linear map (i.e. a matrix)
- $W \subset \mathbb{R}^n$ be an $A$–invariant $k$– dimensional subspace of $\mathbb{R}^n$

There exists a nonsingular $T = (e_1 \ e_2 \ \cdots \ e_n)$ s.t.

- $e_1, \cdots, e_k$ form a basis of $W$,
- in the basis of $e_1, e_2, \cdots, e_n$, $A$ is represented by:

$$T^{-1}AT = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{pmatrix}$$

$\tilde{A}_{11} \in \mathbb{R}^{k \times k}$ where $k = \text{dim } W$. 

M.E. University of Minnesota 42
For any vector \( B \in W \), the representation of \( B \) is given by

\[
B = T\left( \tilde{B} \right).
\]

The point of the theorem is that \( A \) can be put in a form with 0 in strategic locations (\( \tilde{A}_{21} \) are zeros), similarly for any vector \( B \in W \) (\( \tilde{B}_2 \) are zeros).

**Proof:** Take \( e_1, \cdots, e_k \) to be a basis of \( W \), and choose \( e_{k+1}, \cdots e_n \) to complete the basis of \( \mathbb{R}^n \). Because \( Ae_i \in W \) for each \( i = 1, \cdots k \),

\[
Ae_i = \sum_{l=1}^{k} \alpha_{li}e_l.
\]

Therefore, \( \alpha_{li} \) are the coefficients of \( \tilde{A}_{11} \) and the coefficients of \( \tilde{A}_{21} \) are all zero.

Let \( B \in W \), then

\[
B = \sum_{l=1}^{k} \beta_l e_l
\]

so that \( \beta_l \) are simply the coefficients in \( \tilde{B}_1 \) and the coefficients of \( \tilde{B}_2 \) are zeros.
Proof of (4) ⇒ (3) in controllability test: Notice that

\[
A(B \ AB \ A^2B \ \cdots \ A^{n-1}B) = (AB \ A^2B \ \cdots \ A^{n-1}B \ A^nB)
\]

therefore, by the Cayley Hamilton Theorem, the range space of the controllability matrix is \(A\)– invariant.

Suppose that (3) is not true, so that the rank of

\[
\begin{pmatrix}
B & AB & A^2B & \cdots & A^{n-1}B
\end{pmatrix}
\]

is less than \(n\). Since the controllability matrix is \(A\)-invariant, by the 2nd representation theorem, there is an invertible matrix \(T\) so that in the new coordinates, \(z = T^{-1}x\),

\[
\dot{z} = \begin{pmatrix}
\tilde{A}_{11} & \tilde{A}_{12} \\
0 & \tilde{A}_{22}
\end{pmatrix} z + \begin{pmatrix}
\tilde{B} \\
0
\end{pmatrix} u \tag{6}
\]

where \(B = T\begin{pmatrix}
\tilde{B} \\
0
\end{pmatrix}\), and

\[
A = T\begin{pmatrix}
\tilde{A}_{11} & \tilde{A}_{12} \\
0 & \tilde{A}_{22}
\end{pmatrix} T^{-1}.
\]
and the dim of $\tilde{A}_{22}$ is non-zero (it is $n - \text{rank}(C)$).

Let $v_2^T$ be a left eigenvector of $\tilde{A}_{22}$ so that

$$v_2^T(\lambda I - \tilde{A}_{22}) = 0$$

for some $\lambda \in \mathbb{C}$. Define $v^T := (0 \ v_2^T)T^{-1}$. Evaluating $v^T B$, $v^T(\lambda I - A)$ gives

$$v^T B = (0 \ v_2^T)T^{-1}T\begin{pmatrix} \tilde{B} \\ 0 \end{pmatrix} = 0$$

$$v^T(\lambda I - A) = \lambda(0 \ v_2^T)T^{-1}(\lambda TT^{-1} - T\tilde{A}T^{-1})$$

$$= (0 \ v_2^T)(\lambda I - \tilde{A})T^{-1} = 0.$$

This shows that $[\lambda I - A, B]$ has rank less than $n$. 
Remarks:

1. Notice that the controllability matrix is not a function of the time interval. Hence, if a LTI system is controllable over some interval, it is controllable over any (non-zero) interval. *c.f.* with result of linear time varying system.

2. Because of the above fact, we often say that the pair \((A, B)\) is controllable.

3. Controllability test can be done by just examining \(A\) and \(B\) without computing the grammian. The test in (3) is attractive in that it enumerates the vectors in the controllability subspace. However, numerically, since it involves power of \(A\), numerical stability needs to be considered.

4. The PBH Test in (4), or the Haustus test for short, involves simply checking the condition at the eigenvalues. It is because for \((sI - A, B)\) to have rank less than \(n\), \(s\) must be an eigenvalue.

5. The range space of the controllability matrix is of special interests. It is called the **CONTROLLABLE SUBSPACE** and is the set of all states that can
be reached from zero-initial condition. This is $A$-invariant.

6. Using the basis for the **controllable subspace** as part of the basis for $\mathbb{R}^n$, the controllability property can be easily seen in (6).
For discrete time system (3), the conditions for controllability is fairly similar except that we need to consider controllability over period of length greater than $n$.

**Theorem** For the discrete time system (3), the followings are equivalent:

1. The system is controllable over the interval $[0, T]$, with $T \geq n$ (i.e. it can transfer from any arbitrary state at $k = 0$ to any other arbitrary state at $k = T$).

2. The reachability grammian,

$$W_{r,T} := \sum_{l=0}^{T-1} A^l B B^* A^* l$$

is full rank $n$.

3. The controllability matrix

$$(B \ AB \ A^2B \ \cdots \ A^{n-1}B)$$

has rank $n$. Notice that the controllability matrix has dimension $n \times nm$ where $m$ is the dimension of the control vector $u$. 
4. For each \( z \in \mathbb{C} \), the matrix,

\[
(zI - A \ B)
\]

has rank \( n \).

**Proof:** The proof of this theorem is exactly analogous to the continuous time case. However, in the case of the equivalence of (2) and (3), we can notice that

\[
x(k) = A^k x(0) + [B, AB, A^2 B, \cdots A^{n-1} B] \begin{pmatrix} u(k - 1) \\ u(k - 2) \\ \vdots \\ u(0) \end{pmatrix}
\]

so clearly one can transfer to arbitrary states at \( k = n \) iff the controllability matrix is full rank. Controllability does not improve for \( T > n \) is again the consequence of Cayley Hamilton Theorem.
Observability

The observability question asks if the initial state is not given, whether one can determine the state from the input and output. This is equivalent to the question of whether one can determine the initial state $x(0)$ is given the input $u(t), t \in [0, T]$ and the output $y(t), t \in [0, T]$.

**Definition** A state determined dynamical system is called observable on $[t_0, t_1]$ if for all inputs, $u(\tau)$, and outputs $y(\tau), \tau \in [t_0, t_1]$, the state $x(t_0) = x_0$ can be uniquely determined.

For observability, the null space of the observability map on the interval $[t_0, t_1]$, given by:

$$L_o : x_0 \mapsto y(\cdot) = C'(\cdot)\Phi(\cdot, t_0)x_0,$$

$$y(\tau) = C(\tau)\Phi(\tau, t_0)x_0, \quad \forall \tau \in [t_0, t_1]$$

must be trivial (i.e. only contains 0). Otherwise, if $L_o(x_n)(t) = y(t) = 0$, for all $t \in [0, T]$, then for any $\alpha \in \mathbb{R}$,

$$y(t) = C(t)\Phi(t, 0)x = C(t)\Phi(t, 0)(x + \alpha x_n).$$
Recall from the finite rank theorem that if the observability map $L_o$ is a matrix, then the null space of $L_o$, $\mathcal{N}(L_o) = \mathcal{N}(L_o^T L_o)$. For continuous time system, $L_o$ can be thought of as an infinitely long matrix, then the matrix product operation in $L_o^T L_o$ becomes an integral, and is written as $L_o^* L_o$ where $L^*$ denotes the adjoint of the $L$.

The observability Grammian $W_{o,[t_0,t_1]} \in \mathbb{R}^{n \times n}$ is given by:

$$W_{o,[t_0,t_1]} = L_o^* L_o = \int_{t_0}^{t_1} \Phi(t, t_0)^T C^T(t) C(t) \Phi(t, t_0) dt$$

where $L_o^*$ is the adjoint (think transpose) of the observability map.
Least Squares Estimation

Continuous time varying system:

\[ \dot{x}(t) = A(t)x(t) \quad y(t) = C(t)x(t) \]

Observability map on \([0, t_1]\): \(\forall \tau \in [0, t_1]\)

\[ L_o : x(0) \mapsto y(\tau) = C(\tau)\Phi(\tau, 0)x(0) \]

Least squares estimate of \(x(0)\) given \(y(\tau), \tau \in [0, t_1]\).

\[ \hat{x}(0|t_1) = \arg \min_{x_0} \int_0^{t_1} e^T(\tau)e(\tau)d\tau. \]

Solution for matrix case is:

\[ Y = L_o x_0 + E; \]

\[ x^*_o = \text{argmin}_{x_0} E^T E = (L_o^T L_o)^{-1} L_o^T Y \]

The resulting \(E\) satisfies \(E^T L_o = 0\), i.e. \(E\) is normal to \(\text{Range}(L_o)\).
\[ \hat{x}(0|t_1) = W_o^{-1}(t_1) \int_0^{t_1} \Phi^T(\tau, 0)C^T(\tau)y(\tau)d\tau \]

\[ W_o(t_1) = L_o^*L_o = \int_0^{t_1} \Phi^T(\tau, 0)C^T(\tau)C(\tau)\Phi(\tau, 0)d\tau. \]

One can also find estimates for:

\[ \hat{x}(t_2|t_1) = \Phi(t_2, 0)\hat{x}(0|t_1) \]

- \( t_2 < t_1 \): Filtering problem
- \( t_2 > t_1 \): Prediction problem
- \( t_2 = t_1 \): Observer problem

Estimate \( \hat{x}(t|t) \) allows us to effective do state feedback control.
Recursive Least Squares Observer

\[ \hat{x}(t|t) = \Phi(t, 0)W_o^{-1}(t) \int_0^t \Phi^T(\tau, 0)C^T(\tau)y(\tau)d\tau \]

\[ W_o(t) = L_o^*L_o = \int_0^t \Phi^T(\tau, 0)C^T(\tau)C(\tau)\Phi(\tau, 0)d\tau. \]

Find \( \hat{x}(t|t) \) recursively via a differential equation.

\[ \frac{d}{dt}\hat{x}(t|t) = A(t)\Phi(t, 0)W_o^{-1}(t) \int_0^t \Phi^T(\tau, 0)C^T(\tau)y(\tau)d\tau \]

\[ + \Phi(t, 0)\frac{d}{dt}W_o^{-1}(t) \int_0^t \Phi^T(\tau, 0)C^T(\tau)y(\tau)d\tau \]

\[ + \Phi(t, 0)W_o^{-1}(t)\Phi^T(t, 0)C^T(t)y(t) \]

Now by differentiating \( W_o(t)W_o^{-1}(t) = I \), we have:

\[ \frac{d}{dt}W_o^{-1}(t) = -W_o^{-1}(t)\frac{dW_o}{dt}(t)W_o^{-1}(t) \]
where

\[
\frac{d}{dt}W_o(t) = \Phi(t, 0)C^T(t)C(t)\Phi(t, 0)
\]

\[
\frac{d}{dt}\hat{x}(t|t) = A(t)\hat{x}(t|t) - \Phi(t, 0)W^{-1}_o(t)\Phi^T(t, 0)C^T(t)C(t)\hat{x}(t|t) + \Phi(t, 0)W^{-1}_o(t)\Phi^T(t, 0)C^T(t)y(t)
\]

\[
= A(t)\hat{x}(t|t) - \underbrace{\Phi(t, 0)W^{-1}_o(t)\Phi^T(t, 0)C^T(t)\left[C(t)\hat{x}(t|t) - y(t)\right]}_{P(t)}
\]

\[P(t)\] satisfies:

\[
\dot{P}(t) = A(t)P(t) + P(t)A^T(t) - P(t)C^T(t)C(t)P(t)
\]

Initialize \[P(0)\] to be a large non-singular matrix, e.g. \[P(0) = \alpha \cdot I\], where \(\alpha\) is large.

The observer feedback gain is:

\[L(t) = P(t)C^T(t)\]
Covariance Windup

\[ P(t) = \Phi(t, 0) \left[ \int_0^t \Phi^T(\tau, 0) C^T(\tau) C(\tau) \Phi(\tau, 0) d\tau \right]^{-1} \Phi^T(t, 0) \]

\[ = \left[ \int_0^t \Phi^T(\tau, t) C^T(\tau) C(\tau) \Phi(\tau, t) d\tau \right]^{-1} =: K^{-1}(t) \]

The matrix \( K(t) := P^{-1}(t) \) is known as the Information Matrix and satisfies:

\[ \dot{K} = C^T(t) C(t) - A^T(t) K(t) - K(t) A(t) \]

Notice that \( K(t) \) can potentially grow indefinitely so that \( P(t) \) will decrease to 0. In most cases. This means that the observer gain \( L(t) \) will decrease. The observer will asymptotically rely more and more on open loop estimation:

\[ \frac{d}{dt} \hat{x}(t|t) = A(t)\hat{x}(t|t) \]

**Forgetting factor** Forget old information (\( \tau \) far away

M.E. University of Minnesota
from current time): Choose $\lambda > 0$:

$$K(t) = \int_{0}^{t} \Phi^T(\tau, t) C^T(\tau) C(\tau) \Phi(\tau, t) e^{-\lambda(t-\tau)} d\tau;$$

$$\dot{K} = C^T(t) C(t) - A^T(t) K(t) - K(t) A(t) - \lambda K$$

**Covariance reset** If $K(t)$ becomes large, reset it to small number.

- Check maximum singular value of $K(t)$, $\bar{\sigma}K(t)$.
- If $\bar{\sigma}K(t) \geq \lambda_{max}$, $K(t) = \epsilon I$, i.e. $P(t) = I/\epsilon$. 
Observability Tests

The observability question deals with the question of whether one can determine what the initial state $x(0)$ is, given the input $u(t), t \in [0, T]$ and the output $y(t), t \in [0, T]$.

For observability, the null space of the observability map,

$$L_o : x \mapsto y(\cdot) = \Phi(\cdot, 0)x, \quad \tau \in [0, T],$$

must be trivial (i.e. only contains 0). Otherwise, if $L_o(x_n)(t) = y(t) = 0$, for all $t \in [0, T]$, then for any $\alpha \in \mathbb{R}$,

$$y(t) = \Phi(t, 0)x = \Phi(t, 0)(x + \alpha x_n).$$

Just like for controllability, it is inconvenient to check the rank (and the null space) of the $L_o$ which is tall and thin. We can instead check the rank and the null space of the observability grammian given by:

$$W_{o,T} = L_o^* L_o$$

where $L_o^*$ is the adjoint (think transpose) of the observability map.
Proposition \( \text{Null}(L_o) = \text{Null}(L_o^* L_o) \).

Proof:

- Let \( x \in \text{Nul}(L_o) \) so, \( L_o x = 0 \). Then, clearly, \( L_o^T L_o x = L^T 0 \). Therefore, \( \text{Null}(L_o) \subseteq \text{Null}(L_o^* L_o) \).

- Let \( x \in \text{Nul}(L_o^* L_o) \). Then, \( x^T L_o^* L_o x = 0 \). Or, \( (L_o x)^T (L_o x) = 0 \). This can only be true if \( L_o x = 0 \). Therefore, \( \text{Null}(L_o^* L_o) \subseteq \text{Null}(L_o) \).
Observability Tests

**Theorem** For the LTI continuous time system (2), the followings are equivalent:

1. The system is observable over the interval \([0, T]\)

2. The observability grammian,

\[
W_{o,T} := \int_0^T e^{A^*t} C^* C e^{At} dt
\]

is full rank \(n\).

3. The observability matrix \(\begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix}\) has rank \(n\).

Notice that the observability matrix has dimension \(np \times n\) where \(p\) is the dimension of the output vector \(y\).

4. For each \(\lambda \in \mathbb{C}\), the matrix,

\[
\begin{pmatrix} \lambda I - A \\ C \end{pmatrix}
\]
has rank \( n \).

**Proof:** The proof is similar to the ones as in the controllable case and will not be repeated here. Some differences are that instead of considering \( 1 \times n \) vector \( v^T \) multiplying on the left hand sides of the controllability matrix and the grammians, we consider \( n \times 1 \) vector multiplying on the RHS of the observability matrix and the grammian etc.

Also instead of considering the range space of the controllability matrix, we consider the NULL space of the observability matrix.

\[ \text{Remark} \]

1. Again, observability of a LTI system does not depend on the time interval. So, theoretically speaking, if observing the output and input for an arbitrary small amount of time will be sufficient to figure out \( x(0) \). In reality, when more data is available, one can do more averaging to eliminate effects of noise (e.g. using the least square or Kalman Filter approach).

2. The subspace of particular interest is the null space of the controllability matrix. An initial state lying
in this set will generate identically 0 zero-input response. This subspace is called the unobservable subspace.

3. Using the basis of the unobservable subspace as part of the basis of $\mathbb{R}^n$, the observability property can be easily seen.
The tests for observability of LTI discrete time system is given similarly by the following theorem.

**Theorem** For the discrete time LTI system, the followings are equivalent:

1. The system is observable over the interval \([0, T]\), for some \(T \geq n\).

2. The observability grammian,

\[
W_{o,T} := \sum_{k=0}^{T-1} A^k C^* C A^k
\]

is full rank \(n\).

3. The observability matrix

\[
\begin{pmatrix}
C \\
CA \\
CA^2 \\
\vdots \\
CA^{n-1}
\end{pmatrix}
\]

has rank \(n\).

Notice that the observability matrix has dimension \(np \times n\) where \(p\) is the dimension of the output vector \(y\).
4. For each $z \in \mathbb{C}$, the matrix,

\[
(zI - A) \\
C
\]

has rank $n$.

The observability matrix can have the following interpretation: zero-input response is given by:

\[
\begin{pmatrix}
y(0) \\
y(1) \\
y(2) \\
\vdots \\
y(n-1)
\end{pmatrix} =
\begin{pmatrix}
C \\
CA \\
CA^2 \\
\vdots \\
CA^{n-1}
\end{pmatrix} x(0).
\]

Thus, clearly if the observability matrix is of full rank, one can reconstruct $x(0)$ from measurements of $y(0), \ldots, y(n-1)$. 
Kalman Decomposition

Controllable / uncontrollable decomposition

Suppose that the controllability matrix \( C \in \mathbb{R}^{n \times n} \) of a system has rank \( n_1 < n \). Then there exists an invertible transformation, \( T \in \mathbb{R}^{n \times n} \) such that:

\[
    z = T^{-1}x,
\]

\[
    \dot{z} = \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{pmatrix} z + \begin{pmatrix} \tilde{B} \\ 0 \end{pmatrix} u \tag{7}
\]

where \( B = T \begin{pmatrix} \tilde{B} \\ 0 \end{pmatrix} \), and

\[
    A = T \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{pmatrix} T^{-1}.
\]

and the dim of \( \tilde{A}_{22} \) is \( n - n_1 \).
Observable / unobservable decomposition

Hence if the observability matrix is not full rank, then using basis for its null space as the last $k$ basis vectors of $\mathbb{R}^n$, the system can be represented as:

$$
\dot{z} = \begin{pmatrix} \tilde{A}_{11} & 0 \\ \tilde{A}_{21} & \tilde{A}_{22} \end{pmatrix} z + \begin{pmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{pmatrix} u
$$

$$
y = (\tilde{C} \ 0) z
$$

where $C = (\tilde{C} \ 0) T^{-1}$, and

$$
A = T \begin{pmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ 0 & \tilde{A}_{22} \end{pmatrix} T^{-1}.
$$

and the dim of $\tilde{A}_{22}$ is non-zero (and is the dim of the null space of the observability matrix).
Theorem There exists a coordinate transformation $z = T^{-1}x \in \mathbb{R}^n$ such that

$$
\dot{z} = \begin{pmatrix}
\tilde{A}_{11} & 0 & \tilde{A}_{13} & 0 \\
\tilde{A}_{21} & \tilde{A}_{22} & A_{23} & A_{24} \\
0 & 0 & \tilde{A}_{33} & 0 \\
0 & 0 & \tilde{A}_{43} & \tilde{A}_{44}
\end{pmatrix} z + \begin{pmatrix}
\tilde{B}_1 \\
\tilde{B}_2 \\
0 \\
0
\end{pmatrix} u
$$

$$
y = (\tilde{C}_1 \ 0 \ \tilde{C}_2 \ 0) z
$$

Let $T = (t_1 \ t_2 \ t_3 \ t_4)$ (with compatible block sizes), then

- $t_2 \rightarrow (C - \bar{O})$: $t_2 = \text{basis for Range}(C) \cap \text{Null}(O)$.
- $t_1 \rightarrow (C - O)$: $t_1 \cup t_2 = \text{basis for Range}(C)$
- $t_4 \rightarrow (\bar{C} - \bar{O})$: $t_2 \cup t_4 = \text{basis for Null}(O)$.
- $t_3 \rightarrow (\bar{C} - O)$: $t_1 \cup t_2 \cup t_3 \cup t_4 = \text{basis for} \ \mathbb{R}^n$.

Note: Only $t_2$ is uniquely defined.
Stabilizability and Detectability

If the uncontrollable modes $\bar{A}_{33}, \bar{A}_{44}$ are stable (have eigenvalues on the Left Half Plane), then, system is called *stabilizable*.

- One can use feedback to make the system stable;
- Uncontrollable modes decay, so not an issue.

If the unobservable modes $\bar{A}_{22}, \bar{A}_{44}$ are stable (have eigenvalues on the Left Half Plane), then, system is called *detectable*.

- The states $z_2, z_4$ decay to 0
- Eventually, they will have no effect on the output
- State can be reconstructed by ignoring the unobservable states (eventually).
Relation to Transfer Function

For the system

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

The transfer function from \( U(s) \to Y(s) \) is given by:

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

\[
= \frac{C\text{adj}(sI - A)B}{\det(sI - A)} + D.
\]

Note that transfer function is not affected by similarity transform.

From Kalman decomposition, it is obvious that

\[
G(s) = \tilde{C}_1(sI - \tilde{A}_{11})^{-1}\tilde{B}_1 + D
\]

where \( \tilde{A}_{11}, \tilde{B}_1, \tilde{C}_1 \) correspond to the controllable and observable component.
Poles

- they are values of $s \in \mathbb{C}$ s.t. $G(s)$ becomes infinite.
- poles of $G(s)$ are clearly the eigenvalues of $\tilde{A}_{11}$.

Zeros

- For the SISO case, $z \in \mathbb{C}$ (complex numbers) is a zero if it makes $G(z) = 0$.
- Thus, zeros satisfy
  \[
  Cadj(zI - A)B + D\text{det}(zI - A) = 0
  \]
  assuming system is controllable and observable, otherwise, need to apply Kalman decomposition first.
- In the general MIMO case, a zero implies that there can be non-zero inputs $U(s)$ that produce output $Y(s)$ the is identically zero. Therefore, there exists $X(z)$, $U(z)$ such that:
  \[
  zI - A \ X(z) = B \ U(z) \\
  0 = C \ X(z) + D \ U(z)
  \]
This will be true if and only if at \( s = z \)

\[
\text{rank} \begin{pmatrix} sI - A & -B \\ -C & -D \end{pmatrix}
\]

is less than the normal rank of the matrix at other \( s \).

We shall return to the multivariable zeros when we discuss the effect of state feedback on the zero of the system.
Pole-zero cancellation

Cascade system with input/output \((u_1, y_2)\):

• System 1 with input/output \((u_1, y_1)\) has a zero at \(\alpha\) and a pole at \(\beta\).

\[
\dot{x}_1 = A_1 x_1 + B_1 u_1 \\
y_1 = C_1 x_1
\]

• System 2 with input/output \((u_2, y_2)\) has a pole at \(\alpha\) and a zero at \(\beta\).

\[
\dot{x}_2 = A_2 x_2 + B_2 u_2 \\
y_2 = C_2 x_2
\]

• Cascade interconnection: \(u_2 = y_1\).

Then

1. The system pole at \(\beta\) is not observable from \(y_2\)

2. The system pole at \(\alpha\) is not controllable from \(u_1\).
The cascade system is given by:

\[
A = \begin{pmatrix} A_1 & 0 \\ B_2 C_1 & A_2 \end{pmatrix}; \quad B = \begin{pmatrix} B_1 \\ 0 \end{pmatrix}; \quad C = \begin{pmatrix} 0 & C_2 \end{pmatrix}
\]

Consider \( x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \).

- Use PBH test with \( \lambda = \beta \). Does there exist \( x_1 \) and \( x_2 \) s.t.

\[
\begin{pmatrix} \beta I - A_1 & 0 \\ -B_2 C_1 & \beta I - A_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}
\]

- Let \( x_1 \) be eigenvector associated with \( \beta \) for \( A_1 \).

- Let \( x_2 = (\beta I - A_2)^{-1}B_2 C_1 x_1 \) (i.e. solve second row in PBH test).

- Then since

\[
Cx = C_2 x_2 = \left( C_2 (\beta I - A_2)^{-1}B_2 \right) C_1 x_1 = \underbrace{G_2(s=\beta)=0}_{G_2(s=\beta)=0}
\]
and $\beta$ is a zero of system 2, PHB test shows that the mode $\beta$ is not observable.

Similar case for $\alpha$ mode in system 2 not controllable by $u$:

PHB test for controllability for $\alpha$ mode:

\[
\begin{pmatrix}
    x_1^T & x_2^T
\end{pmatrix}
\begin{pmatrix}
    \alpha I - A_1 & 0 & B_1 \\
    -B_2 C_1 & \alpha I - A_2 & 0
\end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}
\]

- Let $x_2^T$ be the left eigenvector of $A_2$ s.t. $x_2^T A_2 = \alpha x_2^T$

- Solve for $x_1^T$ in the 1st column:

\[
x_1^T (\alpha I - A_1) = x_2^T B_2 C_1 \\
\Rightarrow x_1^T = x_2^T B_2 C_1 (\alpha I - A_1)^{-1}
\]

- With this choice of $x_1$ and $x_2$,

\[
x_1^T B_1 = x_2^T B_2 C_1 (\alpha I - A_1)^{-1} B_1 = 0
\]

Thus, by the PHB test, the $\alpha$ mode is not controllable.

Lesson: DO NOT CANCEL OUT UNSTABLE MODES !!!
Degree of controllability / observability

Formal definitions are black and white.

- Degree of controllability and observability can be evaluated by the size of $W_r$ and $W_o$ with infinite time horizon:

\[
W_r(0, \infty) = \lim_{T \to \infty} \int_0^T e^{AT}BB^T e^{AT}e^{AT}dT
\]

\[
= \int_0^\infty e^{At}BB^T e^{At}dt
\]

\[
W_o(0, \infty) = \lim_{T \to \infty} \int_0^T e^{AT}CTCe^{At}dT
\]

- $W_r$ satisfies differential equation:

\[
\frac{d}{dt}W_r(0, t) = AW_r(0, t) + W_r(0, t)A^T + BB^T
\]

- If all eigenvalues of $A$ have strictly negative real parts (i.e. $A$ is stable), then, as $T \to \infty$,

\[
0 = AW_r + W_rA^T + BB^T
\]
• Similarly, if $A$ is stable, as $T \to \infty$,

$$0 = A^T W_o + W_o A + C^T C$$

• Minimum norm control: To reach $x(0) = x_0$ from $x(-\infty) = 0$ in infinite time,

$$u = L_r^T W_r^{-1} x_0$$

$$\min_{u(\cdot)} J(u) = \min_{u(\cdot)} \int_{-\infty}^{0} u^T(\tau) u(\tau) d\tau = x_0^T W_r^{-1} x_0.$$  

• If state $x_0$ is difficult to reach, then $x_0^T W_r^{-1} x_0$ is large $\Rightarrow$ practically uncontrollable.

• Signal in an initial state, if $x(0) = x_0$, then, with $u(\tau) \equiv 0$

$$\int_{0}^{\infty} y^T(\tau) y(\tau) d\tau = x_0^T W_o x_0.$$  

• If $x_0^T W_o x_0$ is small, the signal of $x_0$ in the output $y(\cdot)$ is small, thus, it is hard to observe $\Rightarrow$ practically unobservable.
• Generally, one can look at the smallest eigenvalues $W_r$ and $W_o$, the span of the associated eigenvectors will be difficult to control, or difficult to observe.
**Balanced Realization**

- Reduce the number of states while minimizing effect on I/O response (transfer function)

- State equations from distributed parameters system, P.D.E. via finite elements methods, etc. generate many states.

- If unobservable and uncontrollable $\Rightarrow$ remove.

- What about lightly controllable, lightly observable states?

**Issues:**

- States can be lightly controllable, but heavily observable.

- States can be lightly observable, but heavily controllable.

- Both contribute to significant component to input/output (transfer function) response.
Idea: Exhibit states so that they are simultaneously lightly (heavily) controllable and observable.

Consider a stable system:

\[
\dot{x} = Ax + Bu \\
y = Cx + Du
\]

If \( A \) is stable, controllability and observability grammians can be computed by solving the Lyapunov equations:

\[
0 = AW_r + W_r A^T + BB^T \\
0 = A^T W_o + W_o A + C^T C
\]

- If state \( x_0 \) is difficult to reach, then \( x_0^T W_r^{-1} x_0 \) is large \( \Rightarrow \) practically uncontrollable.

- If \( x_0^T W_o x_0 \) is small, the signal of \( x_0 \) in the output \( y(\cdot) \) is small, thus, it is hard to observe \( \Rightarrow \) practically unobservable.

- Generally, one can look at the smallest eigenvalues \( W_r \) and \( W_o \), the span of the associated eigenvectors will be difficult to control, or difficult to observe.
Transformation of grammians: (beware of which side $T$ goes!)\textsuperscript{1}:

$$z = Tx$$

Minimum norm control should be invariant to coordinate transformation:

$$x_0^T W_{r}^{-1} x_0 = z^T W_{r,z}^{-1} z = x^T T^T W_{r,z}^{-1} T x$$

Therefore

Note correction:

$$W_{r,z}^{-1} = T^T W_{r,z}^{-1} T \Rightarrow W_{r,z} = T \cdot W_{r} T^T.$$  

Similarly, the energy in a state transmitted to output should be invariant:

$$x^T W_o x = z W_{o,z} z \Rightarrow W_{o,z} = T^{-T} W_o T^{-1}.$$  

**Theorem:** (Balanced realization) If the linear time invariant system is controllable and observable, then, there exists an invertible transformation $T \in \mathbb{R}^{n \times n}$ s.t. in the coordinates $z = Tx$, $W_{o,z} = W_{r,z}$.

\textsuperscript{1}You may be more comfortable with $x = Tz$, then just substitute $T$ by $T^{-1}$ in the algorithm below
The balanced realization algorithm:

1. Compute controllability (reachability) and observability grammians:
   >> \( W_r = \text{gram}(\text{sys}, 'c'); \)
   >> \( W_o = \text{gram}(\text{sys}, 'o'); \)

2. Write \( W_o = R^T R \) [One can use SVD]
   
   >> [U,S,V]=svd(W_o);
   >> R = sqrtm(S)*V';

3. Diagonalize \( RW_r R^T \) [via SVD again]:

   \[
   RW_r R^T = U \Sigma^2 U^T,
   \]

   where \( UU^T = I \) and \( \Sigma \) is diagonal.

   >> [U_1,S_1,V_1]=svd(R*W_r*R');
   >> \( \Sigma = \text{sqrtm}(S_1'); \)

4. Take \( T = \Sigma^{-\frac{1}{2}} U^T R \)

   >> \( T = \text{inv}(\text{sqrtm}(\Sigma)) * U_1' * R; \)
5. This gives $W_{r,z} = W_{o,z} = \Sigma$.

\[
\begin{align*}
>> W_{r,z} &= T * W_{r} * T', \\
>> W_{o,z} &= \text{inv}(T)' * W_{o} * \text{inv}(T)
\end{align*}
\]

**Proof:** Direct substitution!

- Once $W_{r,z}$ and $W_{o,z}$ are diagonal, and equal, one can decide to eliminate states that correspond to the smallest few entries.

- Do not remove unstable states from the realization since these states need to be controlled (stabilized)

- It is also possible to maintain some D.C. performance.

Consider the system in the balanced realization form:

\[
\frac{d}{dt} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} + \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} u
\]

\[
y = C_1 z_2 + C_2 z_2 + Du
\]

where $z_1 \in \mathbb{R}^{n_1}$ and $z_2 \in \mathbb{R}^{n_2}$, and $n_1 + n_2 = n$. Suppose that the $z_1$ are associated with
the simultaneously lightly controllable and lightly observable mode.

**Naive truncation - Remove** $z_2$:

\[ \dot{z}_1 = A_{11}z_1 + B_1u \]
\[ y = C_1z_1 + Du \]

However, this does not retain the D.C. (steady state) gain from $u$ to $y$ (which is important from a regulation application point of view).

To maintain the steady state response, instead of eliminating $z_2$ completely, replace $z_2$ by its steady state value:

In steady state:

\[ \dot{z}_2 = 0 = A_{21}z_1 + A_{22}z_2^* + B_2u \]
\[ \Rightarrow z_2^* = -A_{22}^{-1}(A_{21}z_1 + B_2u) \]

This is feasible if $A_{22}$ is invertible ($0$ is not an eigenvalue).
Truncation that maintains steady state gain:

\[
\dot{z}_1 = A_{11}z_1 + A_{12}z_2^* + B_1 u \\
y = C_1 z_1 + C_2 z_2^* + D u
\]

\[
\dot{z}_1 = \begin{bmatrix} A_{11} - A_{12}A^{-1}_{22}A_{21} \end{bmatrix} z_1 + \begin{bmatrix} B_1 - A_{12}A^{-1}_{22}B_2 \end{bmatrix} u \\
y = \begin{bmatrix} C_1 - C_2A^{-1}_{22}A_{21} \end{bmatrix} z_1 + \begin{bmatrix} D - C_2A^{-1}_{22}B_2 \end{bmatrix} u
\]

So that the truncated system \( z_1 \in \mathbb{R}^{n_1} \),

\[
\dot{z}_1 = A_r z_1 + B_r u \\
y = C_r z_1 + D_r u
\]

with have the same steady state response as the original system.

**Question:** If the original system is stable, will be truncated system also stable? It turns out that by truncating the states corresponding to small \( \sigma \)'s, the resulting system is also stable (Glover, 1984).