MCE693/793: Analysis and Control of Nonlinear Systems

Introduction to Nonlinear Controllability and Observability

Hanz Richter
Mechanical Engineering Department
Cleveland State University
Definition of Controllability

Definition: The system

\[ \dot{x} = f(x, u) \]

is (fully) controllable if given initial and final points \( x(t_0) \) and \( x_f \), we can always find an admissible control input \( u(t) \) and a \textit{finite} time \( t_f \) such that

\[ \Phi(x(t_0), t_f) = x_f \]

where \( \Phi \) is the flow of the differential equation

\[ \dot{x} = f(x, u(t)) \]

In plain words, there has to be a control that takes the system to the initial point to the final point in some finite time. Note that there is no steady-state requirement for the final point.
Consider the LTIS

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

with \( m \) inputs and \( n \) states.

Intuitively, the ability to drive the system state from one point to another infinitesimally close to it is related to the possibility of using a constant control input to target desired values for \( \dot{x}, \ddot{x}, \ldots, \frac{d^nx}{dt^n} \) all at once.
Linear Controllability

Setting up equations:

\[
\begin{align*}
\dot{x} &= Ax + Bu \\
\ddot{x} &= A(Ax + Bu) \\
\vdots \\
\frac{dx^n}{dt^n} &= A(A...(Ax + Bu))
\end{align*}
\]

This can be rearranged as

\[
\begin{bmatrix}
\dot{x} - Ax \\
\ddot{x} - A^2x \\
\vdots \\
\frac{dx^n}{dt^n} - A^{n-1}x
\end{bmatrix}
= 
\begin{bmatrix}
B \\
AB \\
\vdots \\
A^{n-1}B
\end{bmatrix}
\]

\[
u = C^T u
\]
Linear Controllability...

In the above problem, \( x \) is the current state (given). \( C \) is the controllability matrix, and its columns

\[
B_1, B_2, ..B_m, (AB)_1...(AB)_m...(A^{n-1}B)_m
\]

must span \( n \)-dimensional space to find a solution for \( u \).

In Matlab, \( C \) is built with » ctrlb(A,B).

Linear systems theory shows that a LTIS is controllable if and only if it is so between \( x_0 = 0 \) and an arbitrary \( x_f \).
Definition of Observability

Definition: The system

\[
\begin{align*}
\dot{x} &= f(x, u) \\
\dot{y} &= h(x)
\end{align*}
\]

is observable if \( y = 0 \) implies \( x = 0 \). A weaker definition is detectability, where \( y = 0 \) implies that \( x \to 0 \) as \( t \to \infty \).

When a system is observable, the initial state can be uniquely determined from \( y(t) \) and \( u(t) \) for \( t \in [t_0, t_f] \), where \( t_f \) is some finite time.
Consider the LTI system with \( m \) inputs, \( p \) outputs and \( n \) states:

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

Again, suppose a constant control is applied from \( x \) (unknown initial point) for an infinitesimal time. We want to find \( x \) by observing \( y \) and its derivatives up to the \( n - 1 \)-th order.
Linear Observability

Setting up equations ($\dot{u} = 0$):

\[
\begin{align*}
\dot{y} &= Cx + Du \\
\dot{y} &= C(Ax + Bu) \\
\vdots \\
\frac{dy^{n-1}}{dt^{n-1}} &= C(A...(Ax + Bu)) = CA^{n-1}x + CBu
\end{align*}
\]

This can be rearranged as

\[
\begin{bmatrix}
y - Du \\
\dot{y} - CBu \\
\vdots \\
\frac{dy^{n-1}}{dt^{n-1}} - CBu
\end{bmatrix}
= 
\begin{bmatrix}
C \\
CA \\
\vdots \\
CA^{n-1}
\end{bmatrix} x = \mathcal{O}x
\]
Linear Observability...

The observability matrix $O$ must be full rank to find a solution for $x$.

In Matlab, $O$ is built with `obsvb(A,B)`.
The Kalman Decomposition

For linear systems, we can always find a linear transformation that reveals the observable/controllable subspaces in the system.
Kalman Controllability Decomposition

Given

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

there is a transformation \( z = Tx \) with \( T \) orthogonal \( (T^{-1} = T') \) such that the transformed system has the form

\[
\begin{align*}
\dot{x} &= \begin{bmatrix}
A_{nc} & 0 \\
A_{21} & A_c
\end{bmatrix} x + \begin{bmatrix}
0 \\
B_c
\end{bmatrix} u \\
y &= [C_{nc} | C_c] x
\end{align*}
\]

In Matlab, use `ctrbf(A,B,C)`.

When \( A_{nc} \) is Hurwtiz, the system is **stabilizable**.
Kalman Observability Decomposition

Given

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

there is a transformation \( z = Tx \) with \( T \) orthogonal \( (T^{-1} = T') \) such that the transformed system has the form

\[
\begin{align*}
\dot{x} &= \begin{bmatrix} A_{no} & A_{12} \\ 0 & A_o \end{bmatrix} x + \begin{bmatrix} B_{no} \\ B_o \end{bmatrix} u \\
y &= [0|C_o]x
\end{align*}
\]

In Matlab, use \texttt{obsvf(A,B,C)}.

When \( A_{no} \) is Hurwitz, the system is \textit{detectable}. 
Lie derivative of a scalar function with respect to a vector field

Let $\mathcal{M}$ be a subset of $\mathbb{R}^n$. Let $f : \mathcal{M} \mapsto \mathbb{R}^n$ be a smooth vector field and let $h : \mathcal{M} \mapsto \mathbb{R}$ a smooth scalar function. The *Lie derivative* of $h$ with respect to $f$, denoted $L_f h$ is the directional derivative in the direction of $f$:

$$L_f h = \nabla h f$$

If $\dot{x} = f(x)$ is a nonlinear system and $V(x)$ is a Lyapunov function candidate, then

$$\dot{V} = L_f V.$$
Lie derivatives as operators

The Lie derivative can be applied recursively:

\[ L_f(L_fh) = \nabla (L_fh)f = L_f^2 h \]

Also, we can use various vector fields:

\[ L_g(L_fh) = \nabla (L_fh)g = L_gL_fh \]

The Lie derivative is not commutative:

For \( h(x, y) = x - y^2 \), \( g(x, y) = [x \ y]^T \) and \( f(x, y) = [y \ x^2]^T \), calculate \( L_gL_fh \) and \( L_fL_gh \)
The Lie Bracket

Let $f$ and $g$ be two smooth vector fields. The Lie bracket of $f$ and $g$ is another vector field defined by

$$[f, g] = \nabla g f - \nabla f g$$

The notation $\text{ad}_f g$ is also used for $[f, g]$ (given a fixed $f$, $\text{ad}_f g$ is the adjoint action of $f$ on the set of all smooth vector fields on $\mathcal{M}$.

The Lie bracket defines a non-associative algebra of smooth vector fields on a manifold). The algebraic properties are:

1. Bilinearity: $[\alpha_1 f_1 + \alpha_2 f_2, g] = \alpha_1 [f_1, g] + \alpha_2 [f_2, g]$
2. Antisymmetry: $[f, g] = -[g, f]$
3. Jacobi identity: $L_{[f,g]} h = L_{\text{ad}_f g} h = L_f L_g h - L_g L_f h$

Combining the given form of bilinearity with antisymmetry shows that

$$[f, \alpha_1 g_1 + \alpha_2 g_2] = \alpha_1 [f_1, g] + \alpha_2 [f_2, g]$$
Recursive Lie Brackets

\[ \text{ad}_{f^2g} = [f, \text{ad}_f g] \]

This can be worked out using the Jacobi identity:

\[ \text{ad}_{f^2g} h = L_{f^2} L_g h - 2L_f L_g L_f h + L_g L_{f^2} h \]

Exercise: Follow the proof of the Jacobi identity in Slotine and Li and use it to verify the above formula.
Back to Linear Controllability

Consider

\[ \dot{x} = Ax + Bu = Ax + B_1u_1 + B_2u_2 + \ldots B_mu_m \]

We revisit the problem of “targeting” desired values for \( \dot{x}, \ddot{x}, \ldots \frac{d^n x}{dt^n} \) simultaneously, using a constant control.

We show that

\[ C = [B_1, \ldots B_m, \text{ad}_f B_1, \ldots \text{ad}_f B_m, \ldots \text{ad}_{f^{n-1}} B_m] \]

Again, \( C \) must have \( n \) linearly independent columns.
Nonlinear Controllability

Consider the class of affine control systems

$$\dot{x} = f(x) + g(x)u$$

where the columns $g_i$ of $g$ span $\mathbb{R}^m$.

Hunt’s theorem (1982): The nonlinear system is (locally) controllable if there exists an index $k$ such that

$$C = [g_1, \ldots g_m, \text{ad}_f g_1, \ldots \text{ad}_f g_m, \ldots \text{ad}_f^k g_1, \ldots \text{ad}_f^k g_m]$$

has $n$ linearly independent columns.

Example (Hunt)

\[
\begin{align*}
\dot{x}_1 &= \cos(\theta)x_3 + \sin(\theta)x_4 \\
\dot{x}_2 &= -\sin(\theta)x_3 + \cos(\theta)x_4 \\
\dot{x}_3 &= u_1 \\
\dot{x}_4 &= u_2
\end{align*}
\]

with \( \theta = \sqrt{x_1^2 + x_2^2} \).

Notes:

1. Controllability is only local. It can be verified near a point, lost away from that point.

2. Linearization does not preserve local controllability properties!
Back to Linear Observability

With $y_i = C_i x = h_i(x)$, we can take successive derivatives of $y$ using the Lie derivative, using vector field $f = A x$:

$$\frac{d^k y_i}{dt^k} = L_f^k h_i$$

Define

$$G = \begin{bmatrix} L_f^0 h_1 & \ldots & L_f^0 h_p \\ \vdots & \ddots & \vdots \\ L_f^{n-1} h_1 & \ldots & L_f^{n-1} h_p \end{bmatrix}$$

In the linear case, $G$ is:

$$G = \begin{bmatrix} C_1 x & \ldots & C_p x \\ \vdots & \ddots & \vdots \\ C_1 A^{n-1} x & \ldots & C_p A^{n-1} x \end{bmatrix}$$

where $C_i$ are the rows of $C$, $i = 1, 2 \ldots p$. 
Linear Observability...

Form a matrix $dG$ with the gradients of the Lie derivatives of $G$:

$$dG = \begin{bmatrix}
    dL^0_f h_1 & \ldots & dL^0_f h_p \\
    \vdots & & \vdots \\
    dL^{n-1}_f h_1 & \ldots & dL^{n-1}_f h_p
\end{bmatrix}$$

In the linear case, $dG$ becomes the observability matrix.
Nonlinear Weak Observability

Theorem (Hermann and Krener, 1977): Let

\[ \dot{x} = f(x, u) \]
\[ \dot{y} = h(x) \]

Let \( G \) be the set of all finite linear combinations formed with the Lie derivatives of \( h_1, h_2, \ldots, h_p \) with respect to \( f \) and constant \( u \). Let \( dG \) denote the set of the gradients of the elements of \( G \).

The system is weakly (locally) observable if \( dG \) contains \( n \) linearly independent vectors.

Example

\[ \begin{align*}
\dot{x}_1 &= \frac{x_1^2}{2} + e^{x_2} + x_2 \\
\dot{x}_2 &= x_1^2 \\
y &= x_1
\end{align*} \]

Note that \( x_2 \) can be found from \( y \) and \( \dot{y} \), and \( x_1 \) can obviously be found from \( y \). Therefore we now that the system is observable. We use the above technique to show weak observability.
Example: Muscle-Driven System

Consider a mass-spring system driven by a Hill muscle model:

where $L_{CEE} + L_{SEE} = L_m = x_1$. Constant $\beta$ is defined by

$$\beta = x_{eq} + \Delta L$$

where $x_{eq}$ is the equilibrium muscle length and $\Delta L$ is the corresponding elongation of the restraining spring of constant $c$. 
Muscle-Driven System...

The dynamics of the system are given by

\[ \dot{x}_1 = x_2 \]
\[ \dot{x}_2 = \frac{1}{m} \left[ -\Phi_S(L_{SEE}) + c(\beta - x_1) \right] \]
\[ \dot{L}_{SEE} = x_2 + u \]

where \( u \) is the contraction speed of the CE, regarded as control input in this simplified example.

\( \Phi_S(L_{SEE}) \) is the force-length relationship for the series elasticity, which contains a deadzone. We use the above results to verify local controllability and weak observability with \( y = x_1 \).

\( u \) is related to the muscle activation \( a \) by an algebraic equation. Controllability analysis is meaningful with \( u \) as the input.