Nonlinear Control Theory

Lecture 8. Observability

Last time

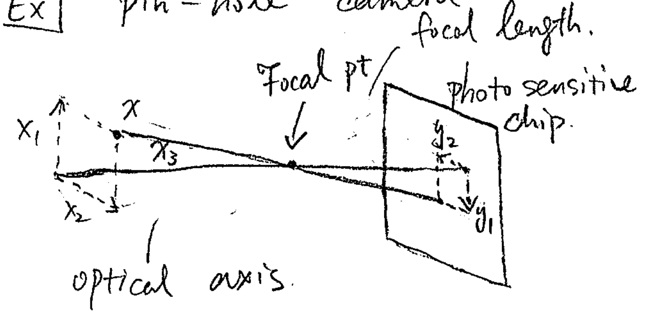
- Controllability
- Lie bracket

Today

- Observability

Nonlinear sensor models

Ex | pin-hole camera



$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ a point's 3D coordinate relative to the focal pt.

The coordinate in the image plane is $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} f \cdot x_1/x_3 \\ f \cdot x_2/x_3 \end{bmatrix}$, $x_3 \neq 0$.

The motion of the point relative to the focal point reads (for $f=1$)

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix},$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1/x_3 \\ x_2/x_3 \end{bmatrix}, \quad x_3 \neq 0.$$

$\omega = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{bmatrix}$ is the angular velocity.

$v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ is the translational velocity.

Observation model is crutial for feedback controller design.

Consider a nonlinear system

$$\dot{x} = f(x,u), \qquad y = h(x).$$

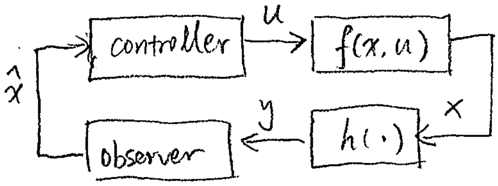
$x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $f, h \in C^1$ in a neighbourhood of the origin, and $f(0,0)=0$, $\forall t$, $h(0)=0$.

Observer design: $\dot{\hat{x}} = \hat{f}(\hat{x}, u, y)$, such that $\|x(t) - \hat{x}(t)\| \to 0$ as $t \to \infty$

Exponential observer: Error $\|x(t) - \hat{x}(t)\|$ converges to zero exponentially fast.

A necessary condition for the existence of an exponential observer is $\left(\frac{\partial f}{\partial x}\Big|_{\substack{x=0 \\ u=0}}, h(0) \right)$ is detectable

Under such condition, locally different choices of input u would barely affect the rate of convergence for an observer. In principle, one can even apply the "separation principle" (the observer & the feedback control law can be designed separately), just as the linear case.



For nonlinear systems, however, its observability does not only depend on the initial value, but also on the control. One "interesting problem" is that how can I design a control to "maximize" the observability?
The topic is called "active sensing", which will not be covered in the course.

Ex (Attitude estimation)

$$\dot{x}_1 = -u_3 x_2 + u_1$$
$$\dot{x}_2 = u_3 x_1 + u_2$$
pitch & roll

$$\dot{x}_3 = \tau x_1 - \tau x_3$$
$$y = x_3$$

Using a low pass sensor to measure the pitch & roll of a rigid body.

One can easily see that, in order to build an observer, u_3 need to satisfy certain constraint.
($u_3 = 0$?)

Consider (*) $\dot{x} = f(x) + g(x)u$, $y = h(x)$, where $x \in M$, $y \in \mathbb{R}^p$, $u \in \mathbb{R}^m$, $f, g, h \in C^1$. $h(0) = 0$ and $x = 0$ is the equilibrium.

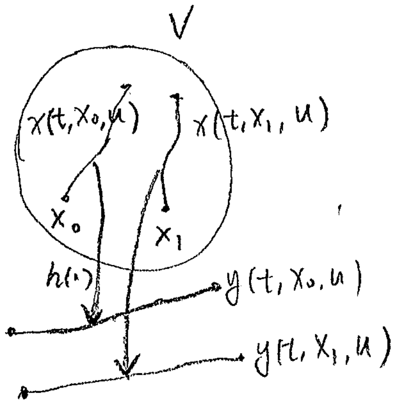
Def. Consider the system (*). Two states x_0 and x_1 are said to be indistinguishable if for all admissible u.
$$y(t, x_0, u) = y(t, x_1, u), \ \forall t$$
where $y(t, x)$ is the output trajectory with initial condition x.

Def. Two states x_0 and x_1 are said to be distinguishable if they are not distinguishable.

Def. Let V be an open set containing x_0 and x_1. x_0 and x_1 are said to be V-distinguishable if there is an admissible control such that

$$y(t,x_0,u) \neq y(t,x_1,u) \quad , t\in[0,T]$$

where $x(t,x_0,u)\in V$ and $x(t,x_1,u)\in V$



find a V, that is all you need.

Def. (observability)

The system is said to be (locally) observable at x_0 if there is a neighbourhood $N(x_0)$ such that every $x\in V\subset N(x_0)$ other than x_0 is V-distinguishable from x_0.

The system is said to be (locally) observable if it is locally observable at every $x\in M$. (Given a control, can different initial value produce the same trajectory?)

Consider linear systems:

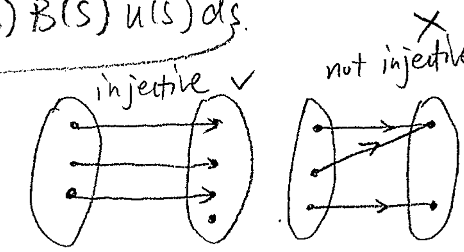
$$\dot{x} = A(t)x + B(t)u$$
$$y = C(t)x + D(t)u$$

$\Rightarrow$ Solution $x(t) = \underbrace{\phi(t,t_0)}_{\text{state transition matrix}} x_0 + \int_0^t \phi(t,s)B(s)u(s)ds$

$$\Rightarrow y(t) = C(t)\phi(t,t_0)x_0 + \int_0^t C(t)\phi(t,s)B(s)u(s)ds + D(t)u(t)$$

$$\Rightarrow \underbrace{C(t)\phi(t,t_0)}_{\Omega(t)}x_0 = \underbrace{y(t) - D(t)u(t) - \int_{t_0}^t C(t)\phi(t,s)B(s)u(s)ds}_{V}$$

It means that $\Omega(t)$ needs to be injective! (one-to-one)



Def. (Lie derivative)

$$L_f h(x) = \frac{\partial h}{\partial x}\cdot f. \quad \text{(The Lie derivative of } h \text{ along } f\text{)}$$

$$\lim_{\alpha\to 0}\frac{h(x+\alpha\cdot f(x)) - h(x)}{\alpha} = \lim_{\alpha\to 0}\frac{\cancel{h(x)} + \alpha\cdot\frac{\partial h}{\partial x}\cdot f(x) + o(\alpha^2) - \cancel{h(x)}}{\alpha} = \frac{\partial h}{\partial x}f(x).$$

Recall the definition of vector fields. It's a mapping assigning to each point $p\in M$ a tangent vector $f(p)\in T_pM$.

As the dual to vector fields, now we study one-forms (co vector fields). Denote T_p^*M the dual space (space of bounded linear functionals) of T_pM, called the cotangent space to M at p. Elements of the cotangent space is called cotangent vectors

The basis of the dual is denoted as $dx_1|_p, \cdots, dx_n|_p$, defined by

$$dx_i|_p \left(\frac{\partial}{\partial x}\right)\Big|_p = \delta_{ij} \quad , i, j = 1, \cdots m$$

$\rightarrow$ kronecker-delta, $\delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$

A one-form ω on M is a mapping assigning to each point $p \in M$ a cotangent vector $\omega(p)$.

Recall that for the system $\begin{aligned} \dot{x} &= f(x) + g(x)u \\ y &= h(x) \end{aligned}$ to be (locally) observable at x^0,

the mapping from x^0 to y must be one-to-one,

$\downarrow$

is affected by $f, g,$ and h.

Let $C^\infty(M)$ denote the infinite dimensional vector space of all C^∞ real valued functions on M. Elements of $\mathfrak{X}(M)$ act as linear operators on $C^\infty(M)$ by ($\mathfrak{X}(M) \subset C^{\infty *}(M)$)

Lie derivative. ($\mathcal{L}_h \varphi(x) = \frac{\partial \varphi}{\partial x} \cdot h = \sum h_i \left(\frac{\partial}{\partial x_i}\right) \varphi$, $h \in \mathfrak{X}(M)$, $\left(\frac{\partial}{\partial x_i}\right)$ is the basis in $\mathfrak{X}(M)$)

Idea Given two initial values x^0 and x^1, if they are V-indistinguishable, then, for any constant admissible controls $u^1, \cdots, u^k \in \mathcal{U}$, small time elapse $s_1, \cdots, s_k \geq 0$ and $h_i, i = 1, \cdots, p$, we have.

$$h_i(\gamma^k_{s_k} \circ \cdots \circ \gamma^2_{s_2} \circ \gamma^1_{s_1}(x^0)) = h_i(\gamma^k_{s_k} \circ \cdots \circ \gamma^2_{s_2} \circ \gamma^1_{s_1}(x^1))$$

flow of $f(x) + g(x)u$!

Differentiating w.r.t s_k yields

$$\mathcal{L}_{X_1} \cdots \mathcal{L}_{X_k} h_i(x^0) = \mathcal{L}_{X_1} \cdots \mathcal{L}_{X_k} h_i(x^1), \quad X_i \in \{f, g_1, \cdots, g_m\}$$

Hence we can define subset $\mathcal{H}^0 = \{h_1, \cdots, h_p\}$ and observable subspace $\mathcal{O}$ as the functions on M with the form $\mathcal{L}_{X_1} \cdots \mathcal{L}_{X_k} h_j$ (it is the smallest linear subspace of $C^\infty(M)$ containing $\mathcal{H}^0$ which is closed w.r.t Lie differentiation by elements of $\{f, g_1, \cdots g_m\}$)

why?

Now, consider a subset of $\mathfrak{X}^*(M)$ by $d\mathcal{H}^0 = \{d\varphi : \varphi \in \mathcal{H}^0\}$, and subspace $d\mathcal{O} = \{d\varphi : \varphi \in \mathcal{O}\}$

Under proper construction of dO (using exact one-form $\omega = d\varphi$), $\mathcal{L}_{X_i}$ and d operation commute, namely, $\mathcal{L}_{X_i}(d\varphi) = d(\mathcal{L}_{X_i}(\varphi))$.

$\Rightarrow$ Elements of dO can be written as

$$d\mathcal{L}_{X_1}\cdots\mathcal{L}_{X_k} h_i = \mathcal{L}_{X_1}\cdots\mathcal{L}_{X_k}(dh_i)$$

Now, if $dO(x^0)$ has full dimension p, then there exists p functions $\varphi_1, \cdots, \varphi_p \in O$, such that $d\varphi_1(x^0), \cdots, d\varphi_p(x^0)$ are linearly independent.

Using <u>inverse function theorem</u>, the mapping $\phi(x^0) = \begin{bmatrix} \varphi_1(x^0) \\ \vdots \\ \varphi_p(x^0) \end{bmatrix}$ is one-to-one.

$\Rightarrow$ observability.

<u>Ex</u> $\dot{x} = Ax + Bu$, $y = Cx$

$$\begin{cases} \mathcal{L}^0_{Ax} Cx = C\cdot x \\ \mathcal{L}^1_{Ax} \cdot Cx = C\cdot Ax \\ \vdots \\ \mathcal{L}^{n-1}_{Ax} \cdot Cx = C\cdot A^{n-1}x \end{cases} \Rightarrow \begin{array}{l} d\mathcal{L}^0_{Ax} Cx = d(C\cdot x) = C \\ d\mathcal{L}^1_{Ax} Cx = d(C\cdot Ax) = CA \\ \vdots \\ d\mathcal{L}^{n-1}_{Ax} Cx = CA^{n-1} \end{array}$$

However, it is known that nonlinear observability does not imply the existence of an observer in general.

$$\begin{aligned} \dot{x}_1 &= -x_1 + x_2^3 \\ \dot{x}_2 &= x_2 + x_1^2 \\ y &= x_1 \end{aligned}$$

this system is observable, but it is shown in

- Xiaoming Hu, On state observers for nonlinear systems, System Control Letters, 17 (1991), No. 6, 465–473

One can not construct any observer for this system.

Hence we focus on some sufficient conditions for observers.

$$\dot{\hat{x}} = p(\hat{x}, h(x(t))), \quad \text{goal}: \|x(t) - \hat{x}(t)\| \to 0 \text{ as } t \to \infty.$$

If the error dynamics is <u>exponentially stable</u>, it's possible to construct an <u>exponential observer</u> of the form $\dot{\hat{x}} = f(\hat{x}) + l(h(x) - h(\hat{x}))$, where $l(0) = 0$. Such observer does not always exist, even if other form of observers exists. (Interested readers can read the above ref.)

Design of exponential observers

Consider $\begin{aligned} \dot{x} &= Ax + f(x) \\ y &= Cx, \end{aligned}$, where $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$. (C, A) detectable

unobservable states are stable.

Linearized system.

Suppose f can be decomposed as $f(x) = f_1(x) + m(Cx)$, where f_1 satisfies a linear growth condition:

$$\|f_1(x) - f_1(z)\| \le k\|x - z\| \ , \forall x, z \in \mathbb{R}^n \quad \text{(global Lipschitz)}$$

Construct the observer.

$$\dot{\hat{x}} = A\hat{x} + f_1(\hat{x}) + m(y) + L(y - C\hat{x})$$

Error dynamics $e = x - \hat{x}$ is

$$\begin{aligned} \dot{e} &= Ax + f_1(x) + m(Cx) - (A\hat{x} + f_1(\hat{x}) + m(\underbrace{y}_{=Cx}) + L(\underbrace{y}_{=Cx} - C\hat{x})) \\ &= (A - LC)e + f_1(x) - f_1(\hat{x}) \end{aligned}$$

<u>Thm.</u> The observer has exponentially stable error dynamics if the solution pair (P, Q) to $(A - LC)^T P + P(A - LC) = -Q$ satisfies $k < \dfrac{\lambda_{min}(Q)}{2\lambda_{max}(P)}$