Lecture 19
Observability and state estimation

• state estimation
• discrete-time observability
• observability – controllability duality
• observers for noiseless case
• continuous-time observability
• least-squares observers
• example
State estimation set up

we consider the discrete-time system

\[ x(t + 1) = Ax(t) + Bu(t) + w(t), \quad y(t) = Cx(t) + Du(t) + v(t) \]

- \( w \) is state disturbance or noise
- \( v \) is sensor noise or error
- \( A, B, C, \) and \( D \) are known
- \( u \) and \( y \) are observed over time interval \([0, t - 1]\)
- \( w \) and \( v \) are not known, but can be described statistically, or assumed small (e.g., in RMS value)
State estimation problem

**state estimation problem:** estimate \( x(s) \) from

\[
u(0), \ldots, u(t-1), y(0), \ldots, y(t-1)\]

- \( s = 0 \): estimate initial state
- \( s = t - 1 \): estimate current state
- \( s = t \): estimate (i.e., predict) next state

an algorithm or system that yields an estimate \( \hat{x}(s) \) is called an *observer* or *state estimator*

\( \hat{x}(s) \) is denoted \( \hat{x}(s|t-1) \) to show what information estimate is based on (read, “\( \hat{x}(s) \) given \( t - 1 \)”)

Observability and state estimation 19–3
Noiseless case

let’s look at finding $x(0)$, with no state or measurement noise:

$$x(t + 1) = Ax(t) + Bu(t), \quad y(t) = Cx(t) + Du(t)$$

with $x(t) \in \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$, $y(t) \in \mathbb{R}^p$

then we have

$$\begin{bmatrix} y(0) \\ \vdots \\ y(t - 1) \end{bmatrix} = O_t x(0) + T_t \begin{bmatrix} u(0) \\ \vdots \\ u(t - 1) \end{bmatrix}$$
where

\[ O_t = \begin{bmatrix} C & CA & \cdots & CA^{t-1} \\ \end{bmatrix}, \quad T_t = \begin{bmatrix} D & 0 & \cdots \\ CB & D & 0 & \cdots \\ & \vdots & \ddots & \ddots & \ddots \\ & & CA^{t-2} B & CA^{t-3} B & \cdots & CB & D \end{bmatrix} \]

- \( O_t \) maps initials state into resulting output over \([0, t-1]\)
- \( T_t \) maps input to output over \([0, t-1]\)

hence we have

\[ O_t x(0) = \begin{bmatrix} y(0) \\ \vdots \\ y(t-1) \end{bmatrix} - T_t \begin{bmatrix} u(0) \\ \vdots \\ u(t-1) \end{bmatrix} \]

RHS is known, \( x(0) \) is to be determined
hence:

- can uniquely determine $x(0)$ if and only if $N(O_t) = \{0\}$

- $N(O_t)$ gives ambiguity in determining $x(0)$

- if $x(0) \in N(O_t)$ and $u = 0$, output is zero over interval $[0, t - 1]$\n
- input $u$ does not affect ability to determine $x(0)$; its effect can be subtracted out
Observability matrix

by C-H theorem, each $A^k$ is linear combination of $A^0, \ldots, A^{n-1}$
hence for $t \geq n$, $\mathcal{N}(O_t) = \mathcal{N}(O)$ where

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

is called the observability matrix

if $x(0)$ can be deduced from $u$ and $y$ over $[0, t - 1]$ for any $t$, then $x(0)$
can be deduced from $u$ and $y$ over $[0, n - 1]$

$\mathcal{N}(O)$ is called unobservable subspace; describes ambiguity in determining
state from input and output

system is called observable if $\mathcal{N}(O) = \{0\}$, i.e., $\text{Rank}(O) = n$
Observability – controllability duality

let \((\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})\) be dual of system \((A, B, C, D)\), i.e.,

\[
\tilde{A} = A^T, \quad \tilde{B} = C^T, \quad \tilde{C} = B^T, \quad \tilde{D} = D^T
\]

controllability matrix of dual system is

\[
\tilde{C} = [\tilde{B} \, \tilde{A}\tilde{B} \cdots \tilde{A}^{n-1}\tilde{B}]
\]

\[
= [C^T \, A^T C^T \cdots (A^T)^{n-1} C^T]
\]

\[
= \mathcal{O}^T,
\]

transpose of observability matrix

similarly we have \(\tilde{O} = C^T\)
thus, system is observable (controllable) if and only if dual system is controllable (observable)

in fact,

\[ \mathcal{N}(O) = \text{range}(O^T)^\perp = \text{range}(\tilde{C})^\perp \]

i.e., unobservable subspace is orthogonal complement of controllable subspace of dual
Observers for noiseless case

suppose $\text{Rank}(O_t) = n$ (i.e., system is observable) and let $F$ be any left inverse of $O_t$, i.e., $FO_t = I$

then we have the observer

$$x(0) = F \left( \begin{bmatrix} y(0) \\ \vdots \\ y(t-1) \end{bmatrix} - T_t \begin{bmatrix} u(0) \\ \vdots \\ u(t-1) \end{bmatrix} \right)$$

which deduces $x(0)$ (exactly) from $u, y$ over $[0, t-1]$

in fact we have

$$x(\tau - t + 1) = F \left( \begin{bmatrix} y(\tau - t + 1) \\ \vdots \\ y(\tau) \end{bmatrix} - T_t \begin{bmatrix} u(\tau - t + 1) \\ \vdots \\ u(\tau) \end{bmatrix} \right)$$
i.e., our observer estimates what state was $t - 1$ epochs ago, given past $t - 1$ inputs & outputs

observer is (multi-input, multi-output) *finite impulse response* (FIR) filter, with inputs $u$ and $y$, and output $\hat{x}$
Invariance of unobservable set

**fact:** the unobservable subspace $\mathcal{N}(\mathcal{O})$ is invariant, i.e., if $z \in \mathcal{N}(\mathcal{O})$, then $Az \in \mathcal{N}(\mathcal{O})$

**proof:** suppose $z \in \mathcal{N}(\mathcal{O})$, i.e., $CA^k z = 0$ for $k = 0, \ldots, n - 1$

evidently $CA^k(Az) = 0$ for $k = 0, \ldots, n - 2$;

$$CA^{n-1}(Az) = CA^n z = - \sum_{i=0}^{n-1} \alpha_i CA^i z = 0$$

(by C-H) where

$$\det(sI - A) = s^n + \alpha_{n-1}s^{n-1} + \cdots + \alpha_0$$
Continuous-time observability

continuous-time system with no sensor or state noise:

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

can we deduce state \( x \) from \( u \) and \( y \)?

let’s look at derivatives of \( y \):

\[
\begin{align*}
y &= Cx + Du \\
\dot{y} &= C\dot{x} + D\dot{u} = CAx + CBu + D\dot{u} \\
\ddot{y} &= CA^2x + CABu + CB\dot{u} + D\ddot{u}
\end{align*}
\]

and so on
hence we have 

\[
\begin{bmatrix}
    y \\
    \dot{y} \\
    \vdots \\
    y^{(n-1)}
\end{bmatrix} = \mathcal{O} x + \mathcal{T}
\begin{bmatrix}
    u \\
    \dot{u} \\
    \vdots \\
    u^{(n-1)}
\end{bmatrix}
\]

where \( \mathcal{O} \) is the observability matrix and 

\[
\mathcal{T} = 
\begin{bmatrix}
D & 0 & \cdots \\
CB & D & 0 & \cdots \\
\vdots & \ddots & \ddots & \ddots \\
CA^{n-2}B & CA^{n-3}B & \cdots & CB & D
\end{bmatrix}
\]

(same matrices we encountered in discrete-time case!)
rewrite as

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

RHS is known; \( x \) is to be determined

hence if \( \mathcal{N}(\mathcal{O}) = \{0\} \) we can deduce \( x(t) \) from derivatives of \( u(t), y(t) \) up to order \( n - 1 \)

in this case we say system is observable

can construct an observer using any left inverse \( F \) of \( \mathcal{O} \):

\[ x = F \left( \begin{bmatrix} y \\ \dot{y} \\ \vdots \\ y^{(n-1)} \end{bmatrix} - T \begin{bmatrix} u \\ \dot{u} \\ \vdots \\ u^{(n-1)} \end{bmatrix} \right) \]
• reconstructs $x(t)$ (exactly and instantaneously) from

$$u(t), \ldots, u^{(n-1)}(t), \ y(t), \ldots, y^{(n-1)}(t)$$

• derivative-based state reconstruction is dual of state transfer using impulsive inputs
A converse

suppose $z \in \mathcal{N}(\mathcal{O})$ (the unobservable subspace), and $u$ is any input, with $x, y$ the corresponding state and output, i.e.,

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

then state trajectory $\tilde{x} = x + e^{tA}z$ satisfies

\[
\dot{\tilde{x}} = A\tilde{x} + Bu, \quad y = C\tilde{x} + Du
\]

i.e., input/output signals $u, y$ consistent with both state trajectories $x, \tilde{x}$

hence if system is unobservable, no signal processing of any kind applied to $u$ and $y$ can deduce $x$

unobservable subspace $\mathcal{N}(\mathcal{O})$ gives fundamental ambiguity in deducing $x$ from $u, y$
Least-squares observers

discrete-time system, with sensor noise:

\[ x(t + 1) = Ax(t) + Bu(t), \quad y(t) = Cx(t) + Du(t) + v(t) \]

we assume \( \text{Rank}(O_t) = n \) (hence, system is observable)

*least-squares* observer uses pseudo-inverse:

\[
\hat{x}(0) = O_t^\dagger \left( \begin{bmatrix} y(0) \\ \vdots \\ y(t-1) \end{bmatrix} - T_t \begin{bmatrix} u(0) \\ \vdots \\ u(t-1) \end{bmatrix} \right)
\]

where \( O_t^\dagger = (O_t^T O_t)^{-1} O_t^T \)
interpretation: $\hat{x}_{ls}(0)$ minimizes discrepancy between

- output $\hat{y}$ that would be observed, with input $u$ and initial state $x(0)$ (and no sensor noise), and

- output $y$ that was observed,

measured as $\sum_{\tau=0}^{t-1} \|\hat{y}(\tau) - y(\tau)\|^2$

can express least-squares initial state estimate as

$$
\hat{x}_{ls}(0) = \left( \sum_{\tau=0}^{t-1} (A^T)^\tau C^T C A^\tau \right)^{-1} \sum_{\tau=0}^{t-1} (A^T)^\tau C^T \tilde{y}(\tau)
$$

where $\tilde{y}$ is observed output with portion due to input subtracted:

$\tilde{y} = y - h \ast u$ where $h$ is impulse response

Least-squares observer uncertainty ellipsoid

since $O_t^\dagger O_t = I$, we have

$$\tilde{x}(0) = \hat{x}_{ls}(0) - x(0) = O_t^\dagger \begin{bmatrix} v(0) \\ \vdots \\ v(t-1) \end{bmatrix}$$

where $\tilde{x}(0)$ is the estimation error of the initial state

in particular, $\hat{x}_{ls}(0) = x(0)$ if sensor noise is zero
(i.e., observer recovers exact state in noiseless case)

now assume sensor noise is unknown, but has RMS value $\leq \alpha$,

$$\frac{1}{t} \sum_{\tau=0}^{t-1} \|v(\tau)\|^2 \leq \alpha^2$$
set of possible estimation errors is ellipsoid

\[ \tilde{x}(0) \in \mathcal{E}_{\text{unc}} = \left\{ \mathcal{O}_t^\dagger \begin{bmatrix} v(0) \\ \vdots \\ v(t-1) \end{bmatrix} \bigg| \frac{1}{t} \sum_{\tau=0}^{t-1} \|v(\tau)\|^2 \leq \alpha^2 \right\} \]

\( \mathcal{E}_{\text{unc}} \) is ‘uncertainty ellipsoid’ for \( x(0) \) (least-square gives best \( \mathcal{E}_{\text{unc}} \))

shape of uncertainty ellipsoid determined by matrix

\[
(\mathcal{O}_t^T \mathcal{O}_t)^{-1} = \left( \sum_{\tau=0}^{t-1} (A^T)\tau C^T CA^\tau \right)^{-1}
\]

maximum norm of error is

\[
\|\hat{x}_{ls}(0) - x(0)\| \leq \alpha \sqrt{t} \|\mathcal{O}_t^\dagger\|
\]
Infinite horizon uncertainty ellipsoid

the matrix

\[ P = \lim_{t \to \infty} \left( \sum_{\tau=0}^{t-1} (A^T)^\tau C^T C A^\tau \right)^{-1} \]

always exists, and gives the limiting uncertainty in estimating \( x(0) \) from \( u, y \) over longer and longer periods:

- if \( A \) is stable, \( P > 0 \)
  \( i.e., \) can't estimate initial state perfectly even with infinite number of measurements \( u(t), y(t), t = 0, \ldots \) (since memory of \( x(0) \) fades . . . )

- if \( A \) is not stable, then \( P \) can have nonzero nullspace
  \( i.e., \) initial state estimation error gets arbitrarily small (at least in some directions) as more and more of signals \( u \) and \( y \) are observed
Example

- particle in $\mathbb{R}^2$ moves with uniform velocity
- (linear, noisy) range measurements from directions $-15^\circ, 0^\circ, 20^\circ, 30^\circ$, once per second
- range noises IID $\mathcal{N}(0, 1)$; can assume RMS value of $v$ is not much more than 2
- no assumptions about initial position & velocity

**problem:** estimate initial position & velocity from range measurements
express as linear system

\[ x(t + 1) = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} x(t), \quad y(t) = \begin{bmatrix} k_1^T \\ \vdots \\ k_4^T \end{bmatrix} x(t) + v(t) \]

- \((x_1(t), x_2(t))\) is position of particle
- \((x_3(t), x_4(t))\) is velocity of particle
- can assume RMS value of \(v\) is around 2
- \(k_i\) is unit vector from sensor \(i\) to origin

true initial position & velocities: \(x(0) = (1 \quad -3 \quad -0.04 \quad 0.03)\)
range measurements (& noiseless versions):

measurements from sensors 1 − 4
• estimate based on \((y(0), \ldots, y(t))\) is \(\hat{x}(0|t)\)

• actual RMS position error is

\[
\sqrt{(\hat{x}_1(0|t) - x_1(0))^2 + (\hat{x}_2(0|t) - x_2(0))^2}
\]

(similarly for actual RMS velocity error)
Observability and state estimation
Continuous-time least-squares state estimation

assume $\dot{x} = Ax + Bu$, $y = Cx + Du + v$ is observable

least-squares estimate of initial state $x(0)$, given $u(\tau)$, $y(\tau)$, $0 \leq \tau \leq t$:
choose $\hat{x}_{ls}(0)$ to minimize integral square residual

$$J = \int_0^t \| \tilde{y}(\tau) - Ce^{\tau A}x(0) \|^2 d\tau$$

where $\tilde{y} = y - h \ast u$ is observed output minus part due to input

let’s expand as $J = x(0)^TQx(0) + 2r^Tx(0) + s,$

$$Q = \int_0^t e^{\tau A^T}C^TCe^{\tau A} d\tau, \quad r = \int_0^t e^{\tau A^T}C^T\tilde{y}(\tau) d\tau, \quad q = \int_0^t \tilde{y}(\tau)^T\tilde{y}(\tau) d\tau$$
setting $\nabla x(0)J$ to zero, we obtain the least-squares observer

$$\hat{x}_{ls}(0) = Q^{-1}r = \left( \int_0^t e^{\tau A^T} C^T C e^{\tau A} d\tau \right)^{-1} \int_0^t e^{A^T \tau} C^T \tilde{y}(\tau) d\tau$$

estimation error is

$$\tilde{x}(0) = \hat{x}_{ls}(0) - x(0) = \left( \int_0^t e^{\tau A^T} C^T C e^{\tau A} d\tau \right)^{-1} \int_0^t e^{A^T \tau} C^T v(\tau) d\tau$$

therefore if $v = 0$ then $\hat{x}_{ls}(0) = x(0)$