ASEN 5044, Fall 2022

Statistical Estimation for Dynamical Systems

Lecture 08:
Intro to Probability

Prof. Nisar Ahmed (Nisar.Ahmed@Colorado.edu)

Thursday 09/15/2022

Ann and H.J. Smead
Aerospace Engineering Sciences
UNIVERSITY OF COLORADO BOULDER

Announcements

- HW 2 due tomorrow Fri 9/10 on Gradescope
- HW 3 posting today, due Fri 09/27
- Quiz 3 + HW 1 solutions posted
- NO QUIZ this week
- In person Lectures resume this Tues 09/24
- **Midterm 1: will be out on Thurs 10/03, due 1 week later (Thurs 10/10)**
  - **Take home exam, to be submitted/graded via Gradescope**
  - **Topic coverage: all lectures + reading + materials for HWs 1-4 + Quizzes 1-4**

Last Time

- Nyquist rate for LTI systems (converting from CT→ DT)
- Stability of LTI CT and DT systems
- Observability in DT LTI Systems

LTI Observability Matrix

Does NOT work for LTV systems generally!

- A **test for LTI observability**: examine rank of **observability matrix**:

$$\mathcal{O} \triangleq \underbrace{\begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-1} \end{bmatrix}}_{(n\cdot p)\times n}$$ = DT LTI observability matrix for system (F,H), i.e.

$$x_{k+1} = Fx_k + Gu_k$$
$$y_k = Hx_k + Mu_k$$

CT LTI: $\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix}$ w/ analagous interpretation

- If rank$(\mathcal{O}) = n$ [# cols = # states], then (F,H) is fully state observable & $R_{\bar{o}} = \{0\}$ [so full col rank]

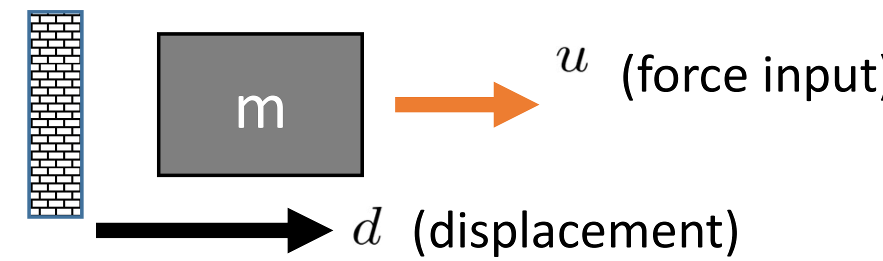
- If rank$(\mathcal{O}) < n$, then (F,H) is not fully state observable & $R_{\bar{o}}$ = Null space$(\mathcal{O})$ [Right null space]

- If the system (F,H) is observable, then it is possible to solve for any x(0) using <u>at most</u> n vector measurements y(0), y(1),…,y(n-1)
- <u>Corollary:</u> if (F,H) observable, then possible to solve for any x(k) by either solving for initial condition x(0) and propagating solution forward, or by using k as initial time and measurements up to no more than y(k+n-1).

Example: 1D Block Mass

- For DT LTI model, we have

$$x(k) = \begin{bmatrix} v(k) \\ d(k) \end{bmatrix} = \begin{bmatrix} \text{velocity at time k} \\ \text{position at time k} \end{bmatrix}$$



$\rightarrow$ Convert from CT to DT (for some fixed Δt sample time):

$$x(k+1) = Fx(k) + Gu(k), \qquad F = \begin{bmatrix} 1 & 0 \\ \Delta t & 1 \end{bmatrix} = \text{STM (see Lec 4)} = \Phi(t, t_0)$$

$$\underbrace{y(k+1)}_{p \times 1,\ p=1} = \underbrace{H}_{p \times n} x(k+1)$$

$\rightarrow$ Suppose $u(k) = 0 \ \forall k \geq 0$ and we have a position sensor: $H = \begin{bmatrix} 0 & 1 \end{bmatrix}$

$\rightarrow$ We get following observability matrix: $\mathbb{O} = \begin{bmatrix} H \\ HF \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \Delta t & 1 \end{bmatrix}$ $\rightarrow$ rank$(\mathbb{O}) = 2$, so system is observable!

$\rightarrow$ But now suppose we instead have a velocity sensor: $H = \begin{bmatrix} 1 & 0 \end{bmatrix} \left(\overbrace{\begin{bmatrix} 1 & 0 \\ \Delta t & 1 \end{bmatrix}}^{F} = \begin{bmatrix} 1 & 0 \end{bmatrix} = HF \right)$

$\rightarrow$ We now get following observability matrix: $\mathbb{O} = \begin{bmatrix} H \\ HF \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$ $\rightarrow$ rank$(\mathbb{O}) = 1$, so system is NOT observable!

Today…

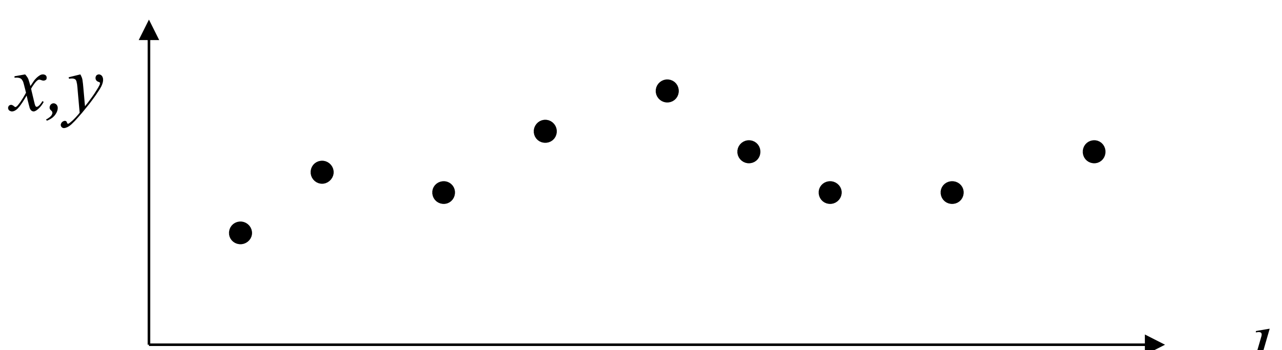
- Intro to Probability
  - Motivation
  - Formal definitions: sample spaces, event spaces, axioms…
  - Basic operations

READ: Chapter 2.3 in Simon book

Intro to Probability

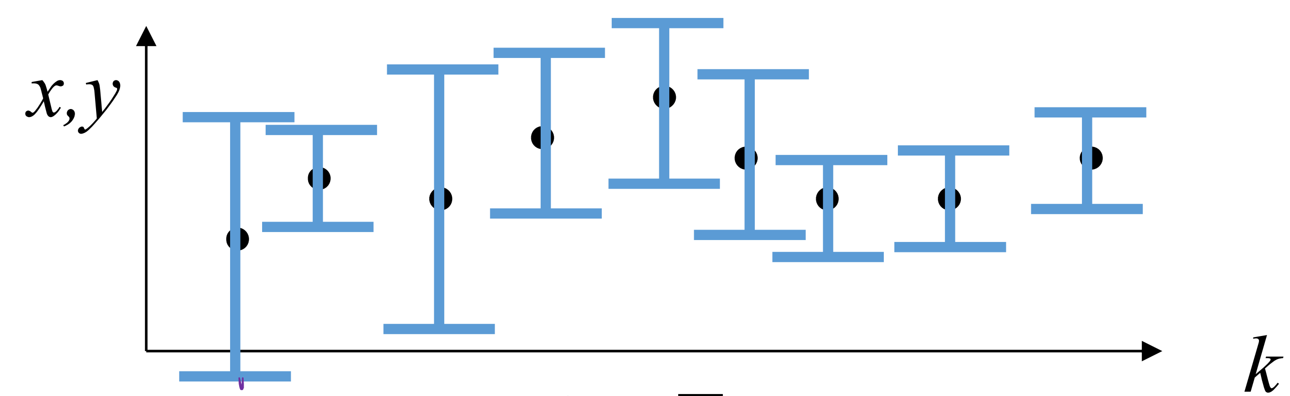
- How to describe **random non-determinism (i.e. random errors/noise)** in SS models?
- Need to use **probability theory**
  - What are probabilities?
  - What are the rules?
  - How to interpret for (scalar/vector-valued) quantities?

(Deterministic variables in DT)



$k \in Z$ (Integers)

(Random non-deterministic variables in DT)



$k \in Z$ (Integers)

Example: 1D Robot Localization with an Inertial Sensor

- Consider robot with inertial position p(t) and acceleration input u(t)
- Robot can measure position with a GPS sensor
- Suppose we take CT model, and discretize with ZOH u(t) and time step Δt

$$x = \begin{bmatrix} p \\ \dot{p} \end{bmatrix} \rightarrow \dot{x} = \begin{bmatrix} \dot{p} \\ \ddot{p} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \longleftrightarrow \text{CT ODE}$$

$$\xrightarrow[\text{ZOH } \Delta t]{\text{DT}} \quad x(k+1) = \begin{bmatrix} p(k+1) \\ \dot{p}(k+1) \end{bmatrix} = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} x(k) + \begin{bmatrix} \frac{1}{2}\Delta t^2 \\ \Delta t \end{bmatrix} u(k)$$

$$y(k+1) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(k)$$

- But robot needs to determine its <u>actual position(k)</u> **and** <u>velocity(k)</u> to drive on road correctly!
- Can position and velocity states at any time step k both be determined if only u and y are available?

$\rightarrow$ check $\mathcal{O}$ [obs. matrix]:

$$\mathcal{O} = \begin{bmatrix} H \\ HF \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & \Delta t \end{bmatrix} \rightarrow \text{rank}(\mathcal{O}) = 2 = \#\text{ states } \checkmark$$

$\rightarrow \therefore$ observable!

Example: 1D Robot Localization with an Inertial Sensor

- But how to account for all of the random physical disturbances that robot encounters?

affects x directly

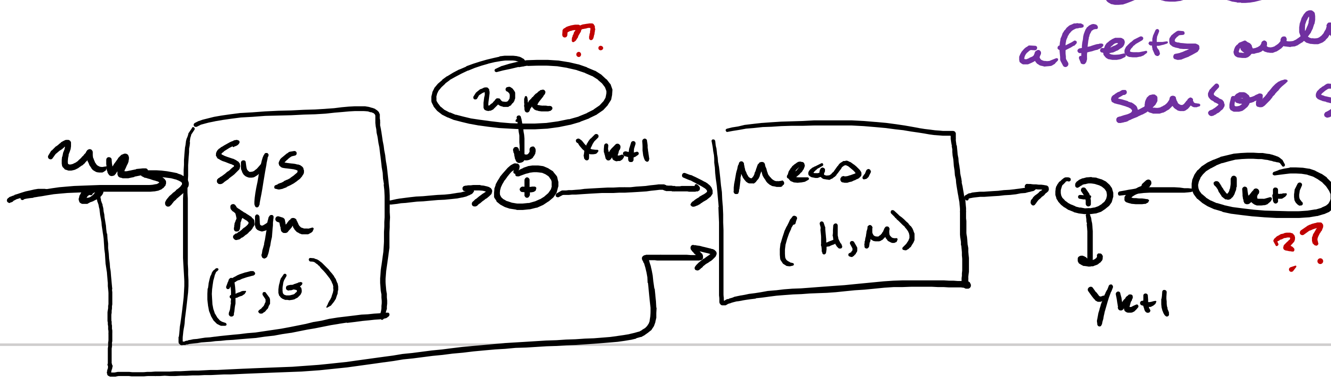
$$x(k+1) = Fx(k) + Gu(k) + \overbrace{w(k)}$$

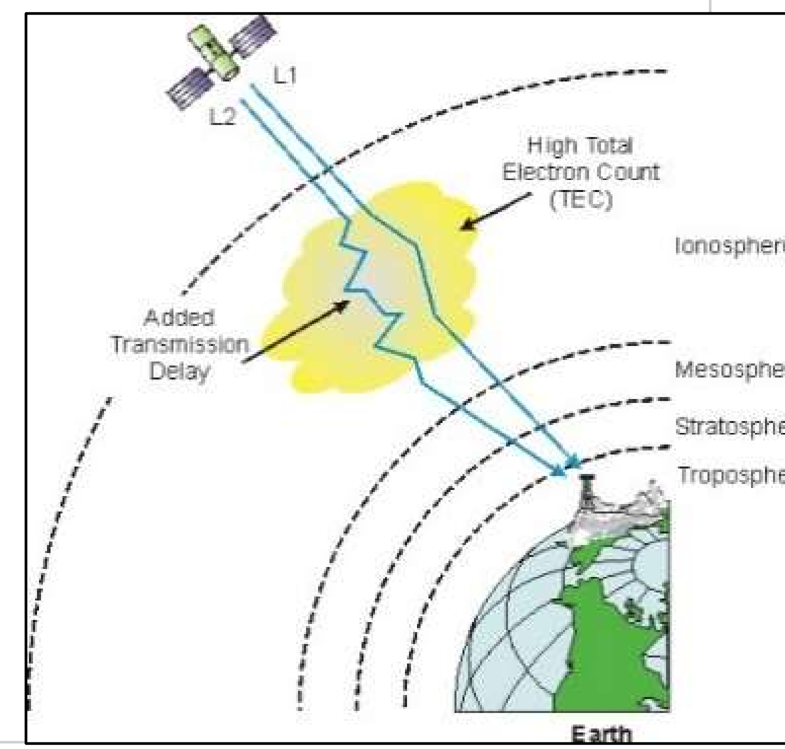
process noise
or
disturbance input

measurement
error

$$y(k+1) = Hx(k+1) + Du(k+1) + \underline{\overline{v(k+1)}}$$

affects only
sensor signals.





Probability as a Description of Randomness

- Probability is a **mathematical tool** for describing **random events, i.e.**
  - **those events which we do not/cannot explicitly know to have occurred for certain…**
  - **…but which would otherwise "explain" why certain other observed events happen**
- "…probability **summarizes uncertainty due to our laziness and ignorance**"
  (Stuart Russell and Peter Norvig, "Artificial Intelligence")

  - Laziness: too much cost/work to predict system behavior exactly (e.g. radio environment)

  - Theoretical ignorance: incomplete domain knowledge (e.g. biomedical; quantum: HUP)

  - Practical ignorance: even if we knew all dynamics, may not know values for all parts
    (e.g. system ID for robot actuators, solar radiation pressure, wind speed, lateral stability derivatives of aircraft)

Probability as a Description of Randomness

- Many different interpretations of probability exist – each has its own issues

(i) "relative frequency": $\text{prob. of event } A = \lim_{N \to \infty} \frac{N_A}{N} = \frac{\#\text{ times event } A \text{ occurs}}{\text{total } \# \text{ experiments}}$

[implicitly assumes repeatability]

→ but in what senses does "$\lim_{N \to \infty}$" exist? (i.e. can't afford to do ∞ # of expts.)

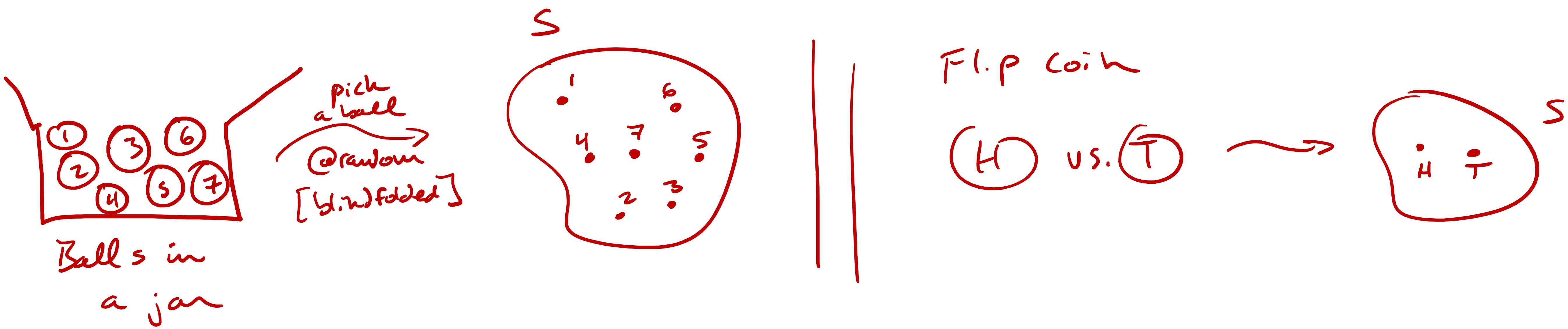
-(ii) "measure of belief" (subjective): make an argument that may or may not be based on experimental data (but they're based on "good old common sense")

Formal Axiomatic Definition of Probability (modern underpinning)

- Define an **experiment** to be a **process with a random outcome**

(i.e. actual outcome is **unpredictable: we can't know what will actually happen,** but the important point is that **we do know what things could possibly happen**)

- **Sample space S**: the set of all possible random outcomes of the experiment

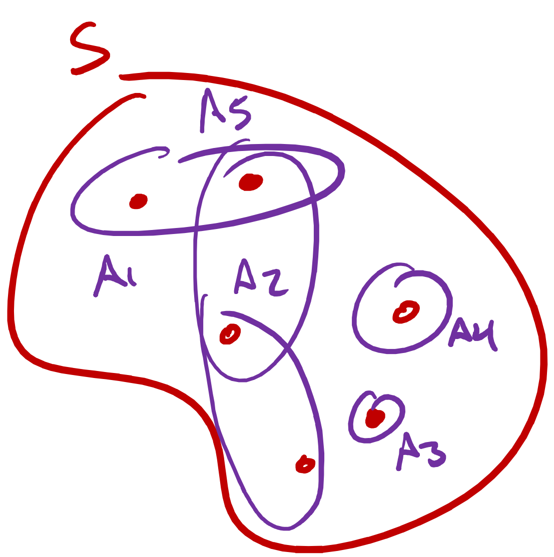


- The sample space S is also sometimes called the **outcome space**

Events and Event Spaces

- **Event**: a subset A of outcomes for the experiment, i.e. A is subset of sample space S
- **Event space**: the set of all events A that are defined on S

- Many ways to define events: each element of S could form its own event, or multiple elements of S could form events, or we can mix and match…



Event i : A_i

Event space : $\{ A_1, A_2, \dots, A_n \}$ (if n events)

- If we observe an outcome in S belonging to event A, then we say that **event A has occurred** (or, **event A has been observed**)

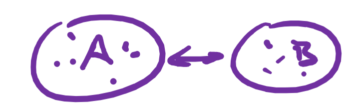
Probabilities and Probability Spaces

- **3 Axioms of Probability**: The **probability of event A**, denoted P(A), satisfies:

(1) $P(A) \geq 0$

(2) $P(S) = 1$

(3) If $A \wedge B$ = set of outcomes in both events A & B & $A \wedge B = \phi$ [A & B don't overlap, ie]

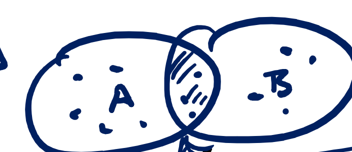


then $P(A \cup B) = P(A) + P(B)$

[$A \cup B$: set of outcomes in either A OR B]

o'wise: $P(A \cup B) = P(A) + P(B) - \underline{P(A \wedge B)}$ [ie $P(A \wedge B) \neq 0$]

[if A & B overlap]



- **Probability space:** a given specification (S, {A_i}, {P(A_i)}) (i.e. a specific sample space S, an event space on S, and probabilities on that event space)
- **This formalism covers both discrete and continuous sample spaces S** (we will focus on discrete case for time being, and move to continuous case later…)

Example: Rolling a Die

- Consider observing the number on top of a standard 6-sided die

$$\text{set of outcomes: } S = \{f_1, f_2, \cdots, f_6\},$$

$$f_i = \text{`outcome: face } i \text{ on top'} \ (i = 1, ..., 6)$$

$Events$: are subsets of S, e.g.

*event $F_1 = \{f_1\}$

*event $F_{even} = \{i \text{ is even}\} = \{f_2, f_4, f_6\}$

*event $F_{\heartsuit} = \{2 < i \leq 4\} = \{f_3, f_4\}$

*sure event: $\{\text{any face i}\} = S$

*Impossible: e.g. $\{i \leq 1 \ \& \ \text{even}\} = \emptyset$

So from axioms:

$$p(F_{even}) = \sum_{i=2,4,6} p(f_i) = \frac{1}{2}$$

$\Rightarrow$ probabilities are *postulated* for each element of S, e.g. $p(f_i) = \frac{1}{6}$ (fair die)