stochastic DE and OCP

goal: mathematical model for stochastic control systems and optimal control problems

---

Markov processes / chains (MP / MC)

AKA stochastic difference equation (SDE)

- the generalization from deterministic to stochastic difference equations is straightforward:

(harder to formalize for differential eqn's; see Higham 2001 for an algorithmic introduction)

- instead of $x^+ = F(x, u)$, consider $x^+ \sim P(x, u)$, where $x \in X$, $u \in \mathcal{U}$, and

$$P : X \times \mathcal{U} \to \Delta(X)$$

$$: (x, u) \mapsto P(x, u)$$

- here, $\Delta(X)$ denotes the set of probability distributions over X, i.e. $\Delta(X) = \{ p : 2^X \to [0,1]$ s.t. p is prob. dist. $\}$,

- so $P(x,u)$ is a function

$$P(x,u) : 2^X \to [0,1]$$

that assigns probability to every event $S \subset X$

(caveats from prior lecture apply: if $|X|, |U| < \infty$ then there's no problem defining P; if $|X|$ or $|U| = \infty$ then need to carefully apply measure theory)

$\longrightarrow$ what is the state of a stochastic DE?

(recall: state contains <u>all</u> and <u>only</u> the information needed to predict the future)

(recall: state contains all and only the information needed to predict the future)

- the state is not a single $x \in X$, but rather a probability distribution $P \in \Delta(X)$

* since this stochastic process has state that evolves in time, it's Markovian:

def: Markov process (MP) specified by $(X, \mathcal{U}, P)$ where $X, \mathcal{U}$ are sets and $P : X \times \mathcal{U} \to \Delta(X)$

- AKA stochastic difference equation (SDE)
- if $\mathcal{U} = \emptyset$ (or $u : X \to \Delta(\mathcal{U})$ given), also termed Markov chain (MC)

Markov decision processes (MDP)

AKA stochastic optimal control problems (SOCP)

- SDE gives natural generalization from a (deterministic) difference equation, so it's natural to formulate analogous stochastic optimal control problem
  - since state is now a distribution over $x \in X$ (and control policy could similarly be distribution over $u \in \mathcal{U}$) we need to summarize cost with a single statistic over ensemble of realizations from controlled SDE

→ propose several statistics that could be used as cost functions; explain when/why they could be useful

- $E[c(x,u)]$ is the average/expected

- $E[c(x,u)]$ is the average/expected cost incurred; useful when system will run many times and/or unlikely outcomes are acceptable (eg HVAC control)

- $\mathrm{Var}[c(x,u)]$ is the variance in cost incurred; useful when system will run few times and/or unlikely outcomes are unacceptable (eg self-driving car)

- we'll focus on minimizing expected cost:

$$\min_u E[c(x,u)] \text{ s.t. } x^+ \sim P(x,u),$$

$$c(x,u) = \ell(t, x_t) + \sum_{s=0}^{t-1} \mathcal{L}(t, x_t, u_t)$$

(finite-horizon)

or $$= \lim \frac{1}{t} \sum^{t} \mathcal{L}(s, x_s, u_s)$$

or $= \lim_{t \to \infty} \frac{1}{t} \sum_{s=0}^{t} \mathcal{L}(s, x_s, u_s)$

(infinite-horizon, average)

or $= \lim_{t \to \infty} \sum_{s=0}^{t} \gamma^s \cdot \mathcal{L}(s, x_s, u_s)$

$\gamma \in (0,1)$ is discount factor

(infinite-horizon, exponentially discounted)

def: Markov decision process (MDP) specified by $(X, \mathcal{U}, P, c)$ where $X, \mathcal{U}$ are sets, $P: X \times \mathcal{U} \to \Delta(X)$, and $c: X^T \times \mathcal{U}^T \to \mathbb{R}$

- T is a time interval, eg $T = [0, t]$ or $T = [0, \infty)$

- also termed stochastic optimal control problem (SOCP)