Markov processes (MP) / stochastic difference equations (SDE)

goal: mathematical model for control systems with random dynamics

refs:

Neuro-Dynamic Programming

Dimitri P. Bertsekas and John N. Tsitsiklis

Chapter 2

Reinforcement Learning

An Introduction

Richard S. Sutton and Andrew G. Barto

Chapter 3

- to generalize from deterministic to stochastic dynamics, we replace $x^+ = F(x,u)$ with $x^+ \sim P(x,u)$

  where $x \in X$, $u \in \mathcal{U}$, $P : X \times \mathcal{U} \to \Delta(X)$

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

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

  ↑ all (measurable) subsets of X

* caveats from prior lectures apply:

- when $|X| < \infty$ there's no problem defining $p \in \Delta(X)$ because p is uniquely determined from $\{p(x)\}_{x \in X}$

- when $|X| = \infty$

because p is uniquely determined from $\{p(\{x\})\}_{x \in X}$

need to be careful about "measurability"

(we'll only consider Gaussians on $X = \mathbb{R}^d$)

→ what is the state of $x^+ \sim P(x, u)$?

- the state is not a single $x \in X$

* instead, the state is a probability distribution $p \in \Delta(X)$

- since this stochastic process has a state that evolves in time, it is referred to as Markovian:

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

- if $\mathcal{U} = \emptyset$ (or if $\underbrace{u : X \to \Delta(\mathcal{U})}_{\text{random / stochastic control policy}}$ is given)

this is termed Markov Chain (MC)