Stochastic Processes and Simple Decisions

Review

Guiding Question

- What does "Markov" mean in "Markov Decision Process"?

Stochastic Process

- A stochastic process is a collection of R.V.s indexed by time.
- $\{x_1, x_2, x_3, \ldots\}$
- $\{x_t\}_{t=1}^{\infty}$ or just $\{x_t\}$

In a *stationary* stochastic process (all in this class), this relationship does not change with time

Example:

$$x_0 = 0$$

$$x_{t+1} = x_t + v_t$$

$v_t \sim \mathcal{U}(\{0, 1\})$ (i.i.d.)

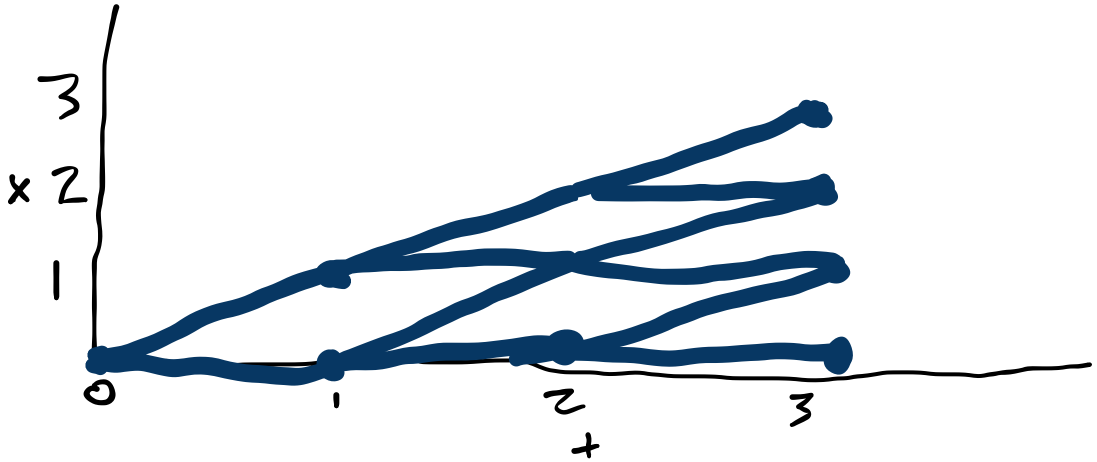
Shorthand:

$$x' = x + v$$

Stochastic Process

$$x_0 = 0 \qquad x_{t+1} = x_t + v_t$$

$$v_t \sim \mathcal{U}(\{0,1\}) \text{ (i.i.d.)}$$



$$P(x_{1:n}) = \prod_{t=1}^{n} P(x_t \mid \text{pa}(x_t))$$

For this particular process,

$$P(x_{1:n}) = \prod_{t=1}^{n} P(x_t \mid x_{t-1})$$

Joint

x0	x1	x2	P(x1, x2, x3)
0	0	0	0.25
0	0	1	0.25
0	1	1	0.25
0	1	2	0.25

Marginal

For this particular process, since $\text{pa}(x_t) = x_{t-1}$, if $P(x_{t-1})$ is known,

$$P(x_t) = \sum_{k \in x_{t-1}} P\left(x_t \mid x_{t-1} = k\right) P(x_{t-1} = k)$$

$$= 0.5\, P(x_{t-1} = x_t - 1) + 0.5\, P(x_{t-1} = x_t)$$

Stochastic Process

Expectation

$$E[x_t] = \sum_{x \in x_t} x P(x_t = x)$$

Expectation of a function
(such as reward)

$$E[f(x_t)] = \sum_{x \in x_t} f(x) P(x_t = x)$$

For this particular process, $x_t = \sum_{i=1}^{t} v_t$, so

$$E[x_t] = E\left[\sum_{i=1}^{t} v_t\right] = \sum_{i=1}^{t} E[v_t] = 0.5t$$

Simulating a Stochastic Process

030-Stochastic-Processes.ipynb

Markov Process

- A stochastic process $\{s_t\}$ is *Markov* if
  $P(s_{t+1} \mid s_t, s_{t-1}, \ldots, s_0) = P(s_{t+1} \mid s_t)$
- s_t is called the "state" of the process

Break

Suppose you want to create a Markov process model that describes how many new COVID cases will start on a particular day. **What information should be in the state of the model?**

Assume:

- The population mixes thoroughly (i.e. there are no geographic considerations).
- COVID patients may be contagious up to 14 days after they contract the disease.
- The number of people infected by each person on day d of their illness is roughly $\mathcal{N}(\mu_d, \sigma^2)$

Hidden Markov Model

(Often you can't measure the whole state)

Simple Decisions

Simple Decisions

Outcomes $S_1 \ldots S_n$

Probabilities $p_1 \ldots p_n$

Lottery $[S_1 : p_1; \ldots ; S_n : p_n]$

- Completeness: Exactly one holds: $A \succ B$, $B \succ A$, $A \sim B$
- Transitivity: If $A \succeq B$ and $B \succeq C$, then $A \succeq C$
- Continuity: If $A \succeq C \succeq B$, then there exists a probability p such that $[A : p; B : 1 - p] \sim C$
- Independence: If $A \succ B$, then for any C and probability p, $[A : p; C : 1 - p] \succeq [B : p; C : 1 - p]$

von Neumann - Morgenstern Axioms

These constraints imply a utility function U with the properties:

- $U(A) > U(B)$ iff $A \succ B$
- $U(A) = U(B)$ iff $A \sim B$
- $U([S_1 : p_1; \ldots ; S_n : p_n]) = \sum_{i=1}^{n} p_i U(S_i)$

Decision Networks

Markov Decision Process

Finite MDP Objectives

1. Finite time

$$\mathrm{E}\left[\sum_{t=0}^{T} r_t\right]$$

2. Average reward

$$\lim_{n\to\infty} \mathrm{E}\left[\sum_{t=0}^{n} r_t\right]$$

3. Discounting

$$\mathrm{E}\left[\sum_{t=0}^{\infty} \gamma^t r_t\right]$$

discount $\gamma \in [0, 1)$

typically 0.9, 0.95, 0.99

if $\underline{r} \leq r_t \leq \bar{r}$

then

$$\frac{\underline{r}}{1-\gamma} \leq \sum_{t=0}^{\infty} \gamma^t r_t \leq \frac{\bar{r}}{1-\gamma}$$

4. Terminal States

Infinite time, but a terminal state (no reward, no leaving) is always reached with probability 1.

Guiding Question

- What does "Markov" mean in "Markov Decision Process"?