Markov Decision Processes **January 16, 2024**

Module 1.B: Random Processes

Lecturer: Avik Kar

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They may be distributed outside this class only with the permission of the Instructor.*

1 Random Processes

Let $\mathcal{T}$ be an arbitrary index set and $x \in \mathbb{R}^{\mathcal{T}}$ be a function. The 'projection operator,' $\boldsymbol{\pi}_t : \mathbb{R}^{\mathcal{T}} \to \mathbb{R}$ maps x to $\boldsymbol{\pi}_t(x) = x_t \in \mathbb{R}$.

Definition 1.1 (Random process)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. For an arbitrary index set $\mathcal{T}$ and state space $\mathcal{X} \subseteq \mathbb{R}$, a map $X : \Omega \to \mathcal{X}^{\mathcal{T}}$ is called a 'random/stochastic process' if the projections $X_t : \Omega \to \mathcal{X}$, defined by $\omega \mapsto X_t(\omega) \triangleq \boldsymbol{\pi}_t(X(\omega))$, are random variables.

Definition 1.2 (Sample path)**.** For each ω, $X(\omega) \in \mathcal{X}^{\mathcal{T}}$ is a function, $X(\omega) : \mathcal{T} \to \mathcal{X}$ called a 'sample path' or a sample function of the process X.

If $\mathcal{T}$ is countable, then we call $X : \Omega \to \mathcal{X}^{\mathcal{T}}$ a *discrete time stochastic process*, else if $\mathcal{T}$ is uncountable, X is a *continuous time stochastic process*.

Definition 1.3 (Distribution of random process)**.** For a random process $X : \Omega \to \mathcal{X}^{\mathcal{T}}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$, we define 'finite dimensional distribution $F_{X_S} : \mathbb{R}^S \to [0,1]$ for a finite $S \subset \mathcal{T}$ by

$$F_{X_S}(x_S) \triangleq \mathbb{P}\left(\prod_{s \in S} X_s^{-1}(-\infty, x_s]\right),\ x_S \in \mathbb{R}^S.$$

Definition 1.4 (Independent random process)**.** A random process $X : \Omega \to \mathbb{R}^{\mathcal{T}}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is called independent if for any finite $S \subseteq \mathcal{T}$,

$$\begin{aligned} F_{X_S}(x_S) &= \prod_{s \in S} \mathbb{P}\left(X_s^{-1}(-\infty, x_s]\right) \\ &= \prod_{s \in S} F_{X_s}(x_s),\ x_S \in \mathbb{R}^S. \end{aligned}$$

Definition 1.5 (Independent and identically distributed random process)**.** A random process $X : \Omega \to \mathbb{R}^{\mathcal{T}}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is said to be independent and identically distributed (*i.i.d.*) with common distribution function $F : \mathbb{R} \to [0,1]$, if for any $s \in \mathcal{T}$,

$$\mathbb{P}\left(X_s^{-1}(-\infty, x]\right) = F(x)\ \forall x \in \mathbb{R}$$

and X is an independent random process, i.e.,

$$\begin{aligned} F_{X_S}(x_S) &= \prod_{s \in S} \mathbb{P}\left(X_s^{-1}(-\infty, x_s]\right) \\ &= \prod_{s \in S} F_X(x_s),\ x_S \in \mathbb{R}^S. \end{aligned}$$

Example 1.6. Consider probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with $\Omega = \{H, T\}^{\mathbb{N}}$, $\mathcal{F} = 2^{\Omega}$. Define $E_n = \{\omega \in \Omega : \omega_n = H\}$ and define

$$\mathbb{P}\left(\cap_{i \in F} E_i\right) \triangleq p^{|F|} \text{ finite } F \subset \mathbb{N}.$$

Then the random sequence $X : \Omega \to \{0,1\}^{\mathbb{N}}$ defined as $X_n(\omega) = \mathbb{1}_{E_n}(\omega)\ \forall \omega \in \Omega$ and $n \in \mathbb{N}$, is *i.i.d.*

Definition 1.7 (Stationary random process)**.** Consider $T \subseteq \mathbb{R}$ is closed under addition. A random process $X : \Omega \to \mathbb{R}^{\mathcal{T}}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is if all finite-dimensional distributions are shift-invariant, that is, for any finite $S \subseteq T$ and $t \in T$, we have

$$F_{X_S}(x_S) = F_{X_{t+S}}(x_S),$$

where $t + S \triangleq \{t + s : s \in S\}$.

Exercise 1.8. Show that a *i.i.d.* process is stationary.

Definition 1.9 (Random walk)**.** Let $X : \Omega \to \mathcal{X}^{\mathbb{N}}$ be an *i.i.d.* random sequence defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathcal{X} \subset \mathbb{R}$ and $\mathbb{E}X_1 \leq \infty$. A random sequence $S : \Omega \to \mathbb{R}^{\mathbb{Z}_+}$ is called a random walk with stepsize sequence X if

$$S_0 \triangleq 0 \text{ and } S_n \triangleq \sum_{i=1}^{n} X_i \text{ for } n \in \mathbb{N}.$$

S is called a simple random walk if $\mathcal{X} \in \{-1, 1\}$.

2 Filtration and stopping times

Let $X : \Omega \to \mathcal{X}^{\mathcal{T}}$ be a random sequance defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Let the index set, $\mathcal{T}$ be ordered.

Definition 2.1 (Filtration)**.** a collection of σ-algebras, $\mathcal{G}_\bullet \triangleq (\mathcal{G}_t \subseteq \mathcal{F} : t \in \mathcal{T})$ is called a filtration if $\mathcal{G}_s \subseteq \mathcal{G}_t\ \forall s, t \in T$ with $s \leq t$.

Definition 2.2 (Natural filtration)**.** The natural filtration associated with a process $X : \Omega \to \mathcal{X}^T$ is given by $\mathcal{F}_\bullet = (\mathcal{F}_t : t \in \mathcal{T})$ where $\mathcal{F}_t \triangleq \sigma(X_s, s \leq t)$.

We say a process X is 'adapted to a filtration' $\mathcal{G}_\bullet$, if $\sigma(X_t) \subseteq \mathcal{G}_t\ \forall t \in \mathcal{T}$. So, a process is always adapted to its natural filtration.

Definition 2.3 (Stopping time)**.** A random variable $\tau : \Omega \to \mathcal{T}$ is called a stopping time with respect to a filtration $\mathcal{G}_\bullet$ if $\tau^{-1}(-\infty, t] \in \mathcal{F}_t\ \forall t \in \mathcal{T}$.

Let $\mathcal{F}_\bullet$ be the natural filtration of a process X. Then 'τ is a stopping time with respect to $\mathcal{F}_\bullet$' means that realization of the event $\{\tau \leq t\}$ can be completely determined by observing $\{X_s : s \leq t\}$.

Exercise 2.4. For a stochastic process $X : \Omega \to \mathcal{X}^{\mathcal{T}}$ and any $A \in \mathcal{B}(\mathcal{X})$, we define the 'first hitting time' to set A, $\tau_X^A : \Omega \to \mathcal{T} \cup \{\infty\}$ as

$$\tau_X^A = \inf\{t \in \mathcal{T} : X_t \in A\}.$$

Show that τ_X^A is a stopping time with respect to the natural filtration of X, $\mathcal{F}_\bullet$.

3 Discrete time Markov chains

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Definition 3.1 (Markov property). A discrete-time random process $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ is said to have the 'Markov property' if for all $n \in \mathbb{Z}_+$

$$\mathbb{P}\left(\{X_{n+1} \le x\} \mid \mathcal{F}_n\right) = \mathbb{P}\left(\{X_{n+1} \le x\} \mid \sigma(X_n)\right),$$

where, $\mathcal{F}_n \triangleq \sigma(X_i, i \le n)$.

Definition 3.2 (Discrete time Markov chain). A stochastic process $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ with countable set $\mathcal{X}$ is called a discrete-time Markov chain (DTMC) if it satisfies the Markov property.

For a DTMC $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$, we have

$$\mathbb{P}\left(\{X_{n+1} = y\} \mid \{X_n = x_n, X_{n-1} = x_{n-1}, \dots, X_0 = x_0\}\right) = \mathbb{P}\left(\{X_{n+1} = y\} \mid \{X_n = x_n\}\right).$$

We define *transition probability kernel*, p as

$$p_{xy}(n) \triangleq \mathbb{P}\left(\{X_{n+1} = y\} \mid \{X_n = x_n\}\right) \text{ for each } n \in \mathbb{Z}_+, x, y \in \mathcal{X}.$$

and define *transition probability matrix* (*t.p.m.*), P such that $P(n) \in [0,1]^{\mathcal{X}\times\mathcal{X}}$ and $P(n)_{x,y} \triangleq p_{xy}(n)$.

Definition 3.3 (Homogeneous DTMC). A DTMC $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ is said to be a homogeneous DTMC if the transition probabilities are time-invariant. So, denote the common transition probability kernel and transition probability matrix as

$$p_{xy} := p_{xy}(n) \text{ and } P := P(n)\ \forall n \in \mathbb{N}.$$

Note that $\sum_{y\in\mathcal{X}} p_{xy} = 1\ \forall x \in \mathcal{X}$, i.e., row sums of of any *t.p.m.* is always 1. Such matrices are called 'stochastic matrices.'

Exercise 3.4. Show that unity is one of the eigenvalue of any stochastic matrix. Then show that for every stochastic matrix P, there exists a distribution, ν with support $\mathcal{X}$, such that

$$\mu = \mu P.$$

Implication: Say $X_n \sim \mu$, i.e., $\mathbb{P}\left(\{X_n = x\}\right) = \mu_x\ \forall x \in \mathcal{X}$. Then,

$$\begin{aligned}\mathbb{P}\left(\{X_{n+1} = y\}\right) &= \sum_{x\in\mathcal{X}} \mathbb{P}\left(\{X_{n+1} = y\} | \{X_n = x\}\right) \mathbb{P}\left(X_n = x\right) \\ &= \sum_{x\in\mathcal{X}} p_{xy}\mu_x = \mu_y\end{aligned}$$

$\implies X_{n+1} \sim \mu$.

Definition 3.5 (Invariant distribution). For a homogeneous DTMC $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ with *t.p.m.* P, a distribution $\mu \in \Delta_{\mathcal{X}}$ is called invariant distribution if it is a left eigenvector of P with eigenvalue unity, i.e.,

$$\mu = \mu P.$$

Definition 3.6 (Transition graph). A homogeneous DTMC $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ with *t.p.m.* P, can be represented by a directed graph $G = (\mathcal{X}, E, w)$, where the set of nodes in G is the state spac, $\mathcal{X}$, the set of edges, E is defined as follows

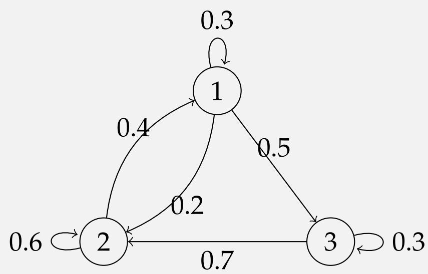
$$E \triangleq \left\{[x, y\rangle \in \mathcal{X} \times \mathcal{X} : p_{xy} > 0\right\},$$

and each edge, $e = [x, y\rangle \in E$ is assigned a weight, $w_e = p_{xy}$.

Example 3.7. Consider the following *t.p.m.* where $\mathcal{X} = [3]$.

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} 0.3 & 0.2 & 0.5 \\ 0.4 & 0.6 & 0 \\ 0 & 0.7 & 0.3 \end{bmatrix}.$$

The corresponding transition graph is



Definition 3.8 (n-step transition probability matrix)**.** For a homogeneous DTMC $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$, we can define n-step transition probability matrix $P^{(n)}$, with its (x, y) entry being the n-step transition probability, denoted by $p_{xy}^{(n)}$ and is defined as

$$p_{xy}^{(n)} = \mathbb{P}\left(\{X_{m+n} = y\} | \{X_m = x\}\right).$$

Theorem 3.9. *The n-step transition probability matrix is given by $P^{(n)} = P^n$ for any $n \in \mathbb{N}$.*

Exercise 3.10. Let $\nu_n \in \Delta_{\mathcal{X}}$ denotes the state distribution at time $n \in \mathbb{Z}_+$, i.e., $\mathbb{P}\left(\{X_n = x\}\right) = \nu_x$. Then show that

$$\nu_{n+1} = \nu_n P, \ \forall n \in \mathbb{Z}_+.$$

Definition 3.11 (Strong Markov Property)**.** Let $\tau : \Omega \to \mathbb{N}$ be a stopping time with respect to a random sequence $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$. Then, for all states $x, y \in \mathcal{X}$ and the event $H_{\tau-1} = \cap_{n=0}\tau - 1\{X_n = x_n\}$, the process X satisfies the strong Markov property if

$$\mathbb{P}\left(\{X_{\tau-1} = y\} | \{X_\tau = x\} \cap H_{n-1}\right) = \mathbb{P}\left(\{X_{\tau-1} = y\} | \{X_\tau = x\}\right)$$

Lemma 3.12. *A homogeneous DTMC satisfies the strong Markov property.*

Theorem 3.13 (Random mapping)**.** *A homogeneous DTMC $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ with finite state space $\mathcal{X}$ has a random mapping representation, i.e., there exists an* i.i.d. *sequence $Z : \Omega \to \mathcal{Z}^{\mathbb{N}}$ and a measurable function $f : \mathcal{X} \times \mathbb{Z} \to \mathcal{X}$ such that for each $n \in \mathbb{N}$*

$$X_n = f(X_{n-1}, Z_n).$$

Definition 3.14 (Communication and communicating classes)**.** Let $x, y \in \mathcal{X}$. If $p_{xy}^{(n)} > 0$ for some $n \in \mathbb{Z}_+$, then we say that y is 'accessible' from x. If y is accessible from x and x is accessible from y, then we say that they are communicating with each other and denoted by $x \leftrightarrow y$. A set of states that communicate with each other is called a 'communicating class.'

Theorem 3.15. *Let $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ be a homogeneous DTMC. Show that $\mathcal{X}$ is partitioned into a set of communicating classes.*

Proof. To prove the theorem, we have to show that

1. Each state belongs to some communicating class.
2. One state can not be in two communicating classes.

If $x \in \mathcal{X}$, then $p_{xx}^{(0)} = 1$, hence $x \leftrightarrow x$. If x does not communicate with any other state, $\{x\}$ is a communicating class by definition. Else if $\exists y \in \mathcal{X}$ such that $x \leftrightarrow y$, then 1 is true too.
Let $x \in \mathcal{C}_1$ and $x \in \mathcal{C}_2$ where $\mathcal{C}_1$ and $\mathcal{C}_2$ are two different communicating classes, and as they are different, $\exists y \in \mathcal{C}_1$ and $z \in \mathcal{C}_2$ such that $y \neq z$. But $\exists n_1, n_2, m_1, m_2 \in \mathbb{Z}_+$ such that

$$p_{xy}^{(n_1)} > 0,\ p_{yx}^{(n_2)} > 0,\ p_{zx}^{(m_1)} > 0 \text{ and } p_{xz}^{(m_2)} > 0.$$

So, we get that

$$p_{zy}^{(n_1+m_1)} \geq p_{zx}^{(m_1)} p_{xy}^{(n_1)} > 0 \text{ and } p_{yz}^{(n_2+m_2)} \geq p_{yx}^{(n_2)} p_{xz}^{(m_2)} > 0$$
$$\implies y \leftrightarrow z. \quad (\Rightarrow\!\Leftarrow)$$

□

Definition 3.16 (Irreducible DTMC)**.** A DTMC with a single communicating class is called 'irreducible.'

Definition 3.17 (Periodicity)**.** For some $x \in \mathcal{X}$, denote

$$A_x := \{n \in \mathbb{N} : p_{xx}(n) > 0\}.$$

The period of state x is defined as

$$d_x \triangleq gcd(A_x). \tag{1}$$

If the period is 1, we say that the state is aperiodic.

Proposition 3.18. *Periodicity is a class property, i.e., if $x, y \in \mathcal{C}$, a communicating class, then $d_x = d_y$.*

Definition 3.19 (Stationary distribution)**.** For a DTMC $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ with *t.p.m.* P, the 'stationary distribution' is defined as

$$\nu_\infty \triangleq \lim_{n\to\infty} \nu_n.$$

Exercise 3.20. Show that for Markov chains with invariant distribution as the initial distribution, its stationary distribution is the initial distribution.

Theorem 3.21. *For a fixed state irreducible and aperiodic Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$, its invariant distribution is unique and is the same as its stationary distribution.*

Definition 3.22. For a homogeneous DTMC $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$, we can define the 'first hitting time' to state $x \in \mathcal{X}$ as

$$\tau_x^+ \triangleq \inf\{n \in \mathbb{N} : X_n = x\}.$$

If $X_0 = x$, then τ_x^+ is called the first return time' to state x.

Show that τ_x^+ is a stopping time.

Definition 3.23 (First hitting time distribution)**.** Denote

$$f_{xy}^{(n)} \triangleq \mathbb{P}\left(\{\tau_x^+ = n\} \middle| \{X_0 = x\}\right)$$

and $f_{xy} \triangleq \sum_{n\in\mathbb{N}} f_{xy}^{(n)}$. The distribution $\left(\left(f_{xy}^{(n)} : n \in \mathbb{N}\right), 1 - f_{xy}\right)$ is called the first hitting time distribution to state y given initial state x.

Definition 3.24 (Transience and recurrence)**.** A state y is said to be 'transient' if $f_{yy} < 1$, 'recurrent' if $f_{yy} = 1$. A recurrent state is called 'positive recurrent' if $\mathbb{E}\left[\tau_y^+ \middle| X_0 = y\right] = \sum_{n\in\mathbb{N}} n f_{yy}^{(n)} < \infty$. Otherwise, a recurrent state is called 'null recurrent.'

Theorem 3.25. *All states in a communicating class are either transient or positive recurrent or null recurrent.*

Theorem 3.26. *For an irreducible Markov chain with finite state space, all states are positive recurrent.*

4 Martingales

Definition 4.1. A random process $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ is called a 'martingale' with respect to a filtration $\mathcal{F}_\bullet$ if

(m1) $\mathbb{E}\left[X_{n+1}|\mathcal{F}_n\right] = X_n \ \forall n \in \mathbb{Z}_+$

(m2) X is $\mathcal{F}_\bullet$-adapted, i.e., $\sigma(X_n) \subseteq \mathcal{F}_n \ \forall n \in \mathbb{N}$.

(m3) X is integrable, i.e., $\mathbb{E}\left|X_n\right| < \infty \ \forall n \in \mathbb{N}$.

If **(m1)** is modified to $\mathbb{E}\left[X_{n+1}|\mathcal{F}_n\right] \geq (\leq) X_n$, then X is called a sub(super)-martingale.

Example 4.2 (Random walk with zero-mean step-size)**.** Let $X : \Omega \to \mathcal{X}^{\mathbb{N}}$ be independent stepsize sequence with $\mathbb{E}X_n = 0 \ \forall n \in \mathbb{N}$. define

$$S_n \triangleq \sum_{i=1}^{n} X_i = S_{n-1} + X_n.$$

Let $\mathcal{F}_\bullet$ be the natural filtration of X. As X is independent,

$$\mathbb{E}\left[X_{n+1}|\mathcal{F}_n\right] = \mathbb{E}X_{n+1} = 0.$$

and as a result,

$$\mathbb{E}\left[S_{n+1}|\mathcal{F}_n\right] = \mathbb{E}\left[S_n|\mathcal{F}_n\right] + \mathbb{E}\left[X_{n+1}|\mathcal{F}_n\right] = S_n.$$

Example 4.3 (Dood's martingale)**.** Consider a random variable $Z : \Omega \to \mathbb{R}$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $\mathcal{F}_\bullet$ be a filtration such that $\mathcal{F}_n \subseteq \mathcal{F} \ \forall n \in \mathbb{N}$. Then, $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ be a martingale with respect to $\mathcal{F}_\bullet$ where $X_n = \mathbb{E}\left[Z|\mathcal{F}_n\right]$.

$$\begin{aligned}\mathbb{E}\left[X_{n+1}|\mathcal{F}_n\right] &= \mathbb{E}\left[\mathbb{E}\left[Z|\mathcal{F}_{n+1}\right]|\mathcal{F}_n\right] \\ &= \mathbb{E}\left[Z|\mathcal{F}_n\right] \\ &= X_n.\end{aligned}$$

Theorem 4.4 (Doob's optimal stopping)**.** *Let $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ be a martingale with respect to filtration $\mathcal{F}_\bullet$ and $\tau : \Omega \to \mathbb{N}$ stopping time with respect to $\mathcal{F}_\bullet$ such that at least one of the following holds.*

a) $\exists\, n \in \mathbb{N}$ *such that* $\tau \leq n$ *with probability* 1.

b) $\mathbb{E}\tau < \infty$, *and for some* $k \in \mathbb{R}_+$ *we have*

$$\sup_{n\in\mathbb{N}} \mathbb{E}\left[|X_n - X_{n-1}||\mathcal{F}_{n-1}\right] < K$$

c) $X_{\tau\wedge n}$ *is uniformly bounded, i.e.,* $\exists\, c \in \mathbb{R}_+$ *such that* $|X_{\tau\wedge n}| \leq c$ *with probability* 1 *for all* $n \in \mathbb{N}$.

Then X_τ is well defined with probability 1, and $\mathbb{E}X_\tau = \mathbb{E}X_0$. Further, when $\{X_n\}_{n\in\mathbb{N}}$ is super/sub-martingale rather than a martingale, equality is replaced by $\leq$ / $\geq$.