This Week's Agenda

Last time

- Exponential Distribution, Poisson Process

Today

- Discrete Time Markov Chain

Discrete-Time Markov Chain

A discrete-time *Markov chain* is

- A discrete-time stochastic process $\{X_0, X_1, \dots\}$
- State space I is countable (i.e. finite or enumerable)
  I is often a subset of $\mathbb{N}$ or $\mathbb{Z}$
- For all $i, j \geq 0$ there is a given probability p_{ij} such that

  $$P\left[X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \dots X_0 = i_0\right] = p_{ij},$$

  for all $i_0, \dots, i_{n-1}, n \geq 0$.
- Clearly, the p_{ij} have to satisfy

  $$p_{ij} \geq 0, \forall i, j \in I, \text{ and } \sum_{j \in I} p_{ij} = 1, \forall i \geq 0.$$
- $\mathbf{P} = (p_{ij})$ is called the *transition probability matrix*

Chapman-Kolmogorov Equations

- A *measure* on I is a vector λ where $\lambda_i \geq 0$, for all $i \in I$.
- A measure is a *distribution* if $\sum_i \lambda_i = 1$.
- For any event F, let $\Pr_i[F] = \Pr[F \mid X_0 = i]$
- For any random variable Z, let $\mathrm{E}_i[Z] = \mathrm{E}[Z \mid X_0 = i]$
- Let $p_{ij}^{(n)} = \Pr_i[X_n = j]$
- Chapman-Kolmogorov Equations:

$$p_{ij}^{(m+n)} = \sum_{k=0}^{\infty} p_{ik}^{(n)} p_{kj}^{(m)}, \forall n, m \geq 0, i, j \in I.$$

- It is clear that $(\mathbf{P}^n)_{ij} = p_{ij}^{(n)}$
- If λ is the distribution of X_0, then $\lambda^T \mathbf{P}^n$ is the distribution of X_n. We also write

$$(X_n)_{n \geq 0} = \mathrm{Markov}(\mathbf{P}, \lambda).$$

Communication Classes

- j is *reachable* from i if $p_{ij}^{(n)} > 0$ for some $n \geq 0$. We write $i \rightsquigarrow j$.
- i and j *communicate* if $i \rightsquigarrow j$ and $j \rightsquigarrow i$. We write $i \leftrightarrow j$.
- Communication is an equivalence relation, partitioning I into *communication classes*
- Communication classes are strongly connected components of the directed graph corresponding to for $\mathbf{P}$
- A chain is irreducible if there is only one class
- A *closed* class C is a class where $i \in C$ and $i \rightsquigarrow j$ imply $j \in C$ (no escape!)
- A state i is *absorbing* if i is a class

Passage Times, Recurrent and Transient States

- *First passage time*:

$$T_i = \inf\{n \geq 1 \mid X_n = i\}. \tag{1}$$

- Define

$$\begin{aligned} f_{ij}^{(n)} &= \Pr_i[X_n = j \ \wedge \ X_s \neq j, \forall s = 1, .., n-1] \\ &= \Pr_i[T_j = n] \\ f_{ij} &= \sum_{n=1}^{\infty} f_{ij}^{(n)} \\ \mu_{ij} &= \sum_{n=1}^{\infty} n f_{ij}^{(n)} \end{aligned}$$

- i is recurrent (or *persistent*) iff $f_{ii} = 1$, and transient if $f_{ii} < 1$.
- A recurrent state is
  - positive if $\mu_{ii} < \infty$
  - null if $\mu_{ii} = \infty$

A Characterization of Recurrence and Transience

Theorem 1. *Given a DTMC* $\mathbf{P}$ *and a state* i*,*

(i) i *is recurrent iff* $\sum_{n\geq 0} p_{ii}^{(n)} = \infty$

(ii) i *is transient iff* $\sum_{n\geq 0} p_{ii}^{(n)} < \infty$

Proof. Let V_i be the number of visits to i, namely

$$V_i := \sum_{n=0}^{\infty} 1_{\{X_n=i\}}.$$

Then,

$$\mathrm{E}_i[V_i] = \sum_{n=1}^{\infty} n f_{ii}^{n-1}(1-f_{ii}) = \frac{1}{1-f_{ii}}.$$

On the other hand,

$$\begin{aligned}
\mathrm{E}_i[V_i] &= \mathrm{E}_i\left[\sum_{n=0}^{\infty} 1_{\{X_n=i\}}\right] \\
&= \sum_{n=0}^{\infty} \mathrm{E}_i[1_{\{X_n=i\}}] \\
&= \sum_{n=0}^{\infty} p_{ii}^{(n)}.
\end{aligned}$$

□

Example of Positive Recurrent States

$$\mathbf{P} = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}, \text{for } 0 < p < 1.$$

Let the states be 0 and 1, then

$$\begin{aligned} f_{00}^{(1)} &= f_{11}^{(1)} = p \\ f_{00}^{(n)} &= f_{11}^{(n)} = (1-p)^2 p^{n-2}, \;\; n \geq 2 \end{aligned}$$

Both states are recurrent. Moreover,

$$\mu_{00} = \mu_{11} = p + \sum_{n=2}^{\infty} n(1-p)^2 p^{n-2} = 2.$$

Hence, both states are positive recurrent states.

Class properties

Theorem 2. *Recurrence, transience, and positive/null recurrent are all class properties.*

Theorem 3. *In a DTMC,*

(i) Every recurrent class is closed

(ii) Every finite, closed class is recurrent

Corollary 4. *In a finite DTMC, all recurrent states are positive recurrent. Moreover, if the chain is also irreducible, then*

(i) all states are positive recurrent.

(ii) $\Pr[T_i < \infty] = 1$
(this is true for irreducible, recurrent chains in general).

Example of Null Recurrent States

Consider a Markov chain where $P_{01} = 1$, and for all $i \geq 1$ we have

$$\begin{aligned} P_{i,i+1} &= \frac{i}{i+1} \\ P_{i,0} &= \frac{1}{i+1}. \end{aligned}$$

Then,

$$\begin{aligned} f_{00}^{(1)} &= 0 \\ f_{00}^{(n)} &= \frac{1}{n(n-1)} \\ \sum_{n=1}^{\infty} f_{00}^{(n)} &= 1 \\ \sum_{n=1}^{\infty} n f_{00}^{(n)} &= \sum_{n=1}^{\infty} \frac{1}{n} = \infty. \end{aligned}$$

Consequently, 0 is a null recurrent state.

Example of Transient States

- Consider a random walk on $\mathbb{Z}$, where $p_{i,i+1} = p$ and $p_{i+1,i} = q$, for all $i \in \mathbb{Z}$, $p + q = 1$, $p, q > 0$.
- The chain is an infinite and closed class.
- For any state i, we have

$$
\begin{aligned}
p_{ii}^{(2n+1)} &= 0 \\
p_{ii}^{(2n)} &= \binom{2n}{n} p^n q^n
\end{aligned}
$$

Hence,

$$
\begin{aligned}
f_{ii} &= \sum_{n=0}^{\infty} p_{ii}^{(2n)} = \sum_{n=0}^{\infty} \binom{2n}{n} p^n q^n \\
&\approx \sum_{n \geq n_0} \frac{\sqrt{4\pi n}(2n/e)^{2n}(1+o(1))}{2\pi n (n/e)^{2n}(1+o(1))} p^n q^n \\
&\approx \sum_{n \geq n_0} \frac{1}{\sqrt{\pi n}} (4pq)^n (1+o(1)).
\end{aligned}
$$

which is ∞ if $p = q$ and finite if $p \neq q$.

- Hence, an infinite closed class could be transient or recurrent.

Brief Summary of Recurrent and Transient Properties

- We often only need to look at closed classes (that's where the chain will eventually end up).
- We can then consider irreducible chains instead.

Let $\mathbf{P}$ be an irreducible chain.

- If $\mathbf{P}$ is finite, then $\mathbf{P}$ is positive recurrent.
- If $\mathbf{P}$ is infinite, then $\mathbf{P}$ could be either transient, or positive recurrent, or null recurrent.

Periodicity, Ergodicity

- For a state $i \in I$, let

$$d_i = \gcd\{n : p_{ii}^{(n)} > 0\}.$$

- When $d_i \geq 2$, state i has *period* d_i.
- When $d_i = 1$, state i is aperiodic.
- A DTMC is periodic if it has a periodic state. Otherwise, the chain is aperiodic.

Theorem 5. *If $i \leftrightarrow j$, then $d_i = d_j$. In particular, aperiodicity and periodicity are class properties.*

Theorem 6. *If i is aperiodic, then $\exists n_0 : p_{ii}^{(n)} > 0$, $\forall n \geq n_0$.*

Corollary 7. *If $\mathbf{P}$ is irreducible and has an aperiodic state i, then $\mathbf{P}^n$ has all strictly positive entries for sufficiently large n.*

- An ergodic state is an aperiodic, positive recurrent state.
- An *ergodic Markov chain* is a Markov chain in which all states are ergodic. (Basically, a "well-behaved" chain.)

Stationary Distribution

- A distribution λ is a *stationary* (*equilibrium*, *invariant*) distribution if $\lambda^T \mathbf{P} = \lambda$

Theorem 8. *We have*

(i) Let $(X_n)_{n\geq 0} = \text{Markov}(\mathbf{P}, \lambda)$*, where* λ *is stationary, then* $(X_{n+m})_{n\geq 0} = \text{Markov}(\mathbf{P}, \lambda)$ *for any fixed* m*.*

(ii) In a finite DTMC, suppose for some $i \in I$ *we have*

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j, \forall j \in I,$$

then $\pi = (\pi_j : j \in I)$ *is an invariant distribution.*

- In an infinite DTMC, it is possible that $\lim_{n\to\infty} p_{ij}^{(n)}$ exists for all i, j, producing a vector π for each i, yet π is not a distribution.

- Consider the DTMC with state space $\mathbb{Z}$ and

$$p_{i,i+1} = p = 1 - q = 1 - p_{i,i-1}, \;\; \forall i \in \mathbb{Z}.$$

It is not difficult, although tedious, to show that

$$\lim_{n\to\infty} p_{ij}^{(n)} = 0, \forall i, j.$$

Existence of a Stationary Distribution

Theorem 9. *An irreducible DTMC* $\mathbf{P}$ *has a stationary distribution if and only if one of its states is positive recurrent. Moreover, if* $\mathbf{P}$ *has a stationary distribution* π*, then* $\pi_i = 1/\mu ii$.

Two other questions:

- When does an irreducible DMTC converge to a unique invariant distribution?
- What's the long-term behavior of the chain, e.g. how often it visits a state?

Convergence to equilibrium

Theorem 10. *Suppose* $\mathbf{P}$ *is irreducible and ergodic. Then, it has an invariant distribution* π. *Moreover,*

$$\frac{1}{\mu_{jj}} = \pi_j = \lim_{n\to\infty} p_{ij}^{(n)}, \ \forall j \in I.$$

Thus, π *is the unique invariant distribution of* $\mathbf{P}$.

Note: there is a generalized version of this theorem for irreducible chains with period $d \geq 2$. (And the chain is not even required to be positive recurrent.)

Ergodic Theorem

Let

$$V_i(n) = \sum_{k=0}^{n-1} 1_{\{X_k = i\}}.$$

Theorem 11 (Ergodic Theorem). *Let* $\mathbf{P}$ *be an irreducible DTMC. Then*

$$\Pr\left[\lim_{n\to\infty} \frac{V_i(n)}{n} = \frac{1}{\mu_{ii}}\right] = 1$$

Moreover, if $\mathbf{P}$ *is positive recurrent, implying* $\mathbf{P}$ *has a unique invariant distribution* π*, then for any bounded function* $f : I \to \mathbb{R}$,

$$\Pr\left[\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(X_k) = \bar{f}\right] = 1,$$

where

$$\bar{f} = \sum_{i\in I} \pi_i f_i.$$

- Note that, in the former case $f(X_k) = 1_{\{X_k = i\}}$.
- This is essentially the strong law of large numbers for DTMC.