STATS 217: Introduction to Stochastic Processes I

Lecture 21

Stationary times

- Last time, we used the following description of the lazy random walk on the hypercube: suppose the current state is x and the current time is t. We choose a coordinate $i_{t+1} \in \{1, \ldots, n\}$ uniformly at random and an unbiased bit $b_{t+1} \in \{0, 1\}$, also uniformly at random, and independently of the coordinate i_{t+1}. Then, we set the value of coordinate i_{t+1} to b_{t+1} and keep all other coordinates unchanged.
- Let τ_{refresh} be the first time that all coordinates have been chosen to be updated.
- Note that τ_{refresh} is a stopping time with respect to the collection of random tuples $\{(i_k, b_k)\}_{k \geq 1}$. (Why?)
- Moreover, note that $X_{\tau_{\text{refresh}}}$ is distributed uniformly on $\{0, 1\}^n$. (Why?)

Strong stationary times

- τ_{refresh} is an example of a **stationary time**.
- By this, we mean a stopping time τ (with respect to the initial state x of the chain as well as any auxiliary randomness) which satisfies

$$\mathbb{P}[X_\tau = y \mid X_0 = x] = \pi(y) \quad \forall y \in S,$$

  where π is the stationary distribution of the chain.
- Let $x \in S$ and consider the chain with initial state $X_0 = x$. A stationary time τ for which τ and X_τ are independent i.e. for all integers t and $y \in S$

$$\mathbb{P}[\tau = t, X_\tau = y \mid X_0 = x] = \mathbb{P}[\tau = t \mid X_0 = x] \cdot \pi(y)$$

  is called a **strong stationary time** for the starting state x.
- τ_{refresh} is an example of a strong stationary time for any starting state x. (Why?)

Strong stationary times

- A strong stationary time for the starting state x is a stopping time (with respect to the chain and auxiliary randomness) satisfying

$$\mathbb{P}[\tau = t, X_\tau = y \mid X_0 = x] = \mathbb{P}[\tau = t \mid X_0 = x] \cdot \pi(y).$$

- The property that will be useful to us is

$$\mathbb{P}[\tau \leq t, X_t = y \mid X_0 = x] = \mathbb{P}[\tau \leq t \mid X_0 = x] \cdot \pi(y),$$

which can be proved using the law of total probability.

Strong stationary times

Indeed,

$$
\begin{aligned}
\mathbb{P}[\tau \leq t, X_t = y \mid X_0 = x] &= \sum_{s \leq t} \sum_{z \in S} \mathbb{P}[\tau = s, X_t = y, X_s = z \mid X_0 = x] \\
&= \sum_{s \leq t} \sum_{z \in S} \mathbb{P}[\tau = s, X_s = z \mid X_0 = x] \cdot P_{z,y} \\
&= \sum_{s \leq t} \sum_{z \in S} \left(\mathbb{P}[\tau = s \mid X_0 = x] \cdot \pi(z) \right) P_{z,y} \\
&= \sum_{s \leq t} \mathbb{P}[\tau = s \mid X_0 = x] \cdot \sum_{z \in S} \pi(z) P_{z,y} \\
&= \sum_{s \leq t} \mathbb{P}[\tau = s \mid X_0 = x] \cdot \pi(y) \\
&= \mathbb{P}[\tau \leq t \mid X_0 = x] \cdot \pi(y).
\end{aligned}
$$

Bounding the mixing time using strong stationary times

- Strong stationary times are also a useful tool for bounding the mixing time. This is captured by the following:
- Suppose that τ is a strong stationary time for the starting state x. Then,

$$\mathrm{TV}(X_t \mid X_0 = x, \pi) \le \mathbb{P}[\tau > t \mid X_0 = x].$$

- Note that

$$\begin{aligned}
\mathrm{TV}(X_t \mid X_0 = x, \pi) &= \sum_{y:\pi(y) > P^t_{x,y}} [\pi(y) - P^t_{x,y}] \\
&= \sum_{y:\pi(y) > P^t_{x,y}} \pi(y) \left(1 - \frac{P^t_{x,y}}{\pi(y)}\right) \\
&\le \max_{y \in S} \left(1 - \frac{P^t_{x,y}}{\pi(y)}\right).
\end{aligned}$$

Bounding the mixing time using strong stationary times

- Hence, it will suffice to show that for all $y \in S$,

$$1 - \frac{P^t_{x,y}}{\pi(y)} \leq \mathbb{P}[\tau > t \mid X_0 = x].$$

- We have

$$\begin{aligned} 1 - \frac{P^t_{x,y}}{\pi(y)} &= 1 - \frac{\mathbb{P}[X_t = y \mid X_0 = x]}{\pi(y)} \\ &\leq 1 - \frac{\mathbb{P}[X_t = y, \tau \leq t \mid X_0 = x]}{\pi(y)} \\ &= 1 - \frac{\pi(y)\mathbb{P}[\tau \leq t \mid X_0 = x]}{\pi(y)} \\ &= \mathbb{P}[\tau > t \mid X_0 = x]. \end{aligned}$$

Example: top-to-random shuffle

- Consider a deck of n cards.
- At each step of the top-to-random shuffle, we take the top card, and place it in any of the n available positions.
- On the homework, you showed that this is an aperiodic and irreducible chain with the unique stationary distribution π given by the uniform distribution on all possible permutations of the n cards.
- Now, we will bound the mixing time by constructing a suitable strong stationary time.

Example: top-to-random shuffle

- Let τ_{top} be first time when the original bottom card has moved to the top.
- Then, $\tau = \tau_{\text{top}} + 1$ is a strong stationary time.
- Clearly, this is a stopping time.
- The key to showing that it is a strong stationary time is the following observation, which can be proved by induction: let τ_k denote the first time when there are k cards under the original bottom card. Then, at X_{τ_k}, all $k!$ permutations of these cards are equally likely. (Why?)
- Therefore, by the relationship between strong stationary times and total variation distance,

$$\text{TV}(X_t \mid X_0 = x, \pi) \leq \mathbb{P}[\tau > t \mid X_0 = x].$$

Example: top-to-random shuffle

- Note that, when there are k cards under the bottom card, the probability that the top card goes under it is $(k+1)/n$.

- Therefore, we see that

$$\tau \sim \text{Geom}(1/n) + \text{Geom}(2/n) + \cdots + \text{Geom}(1),$$

  where the geometric random variables on the right hand side are independent. This is the same as the distribution of the coupon collector's time.

- Therefore, as for the lazy random walk on the hypercube, we get that

$$\tau_{\text{mix}}(\varepsilon) \leq n \log n + n \log(\varepsilon^{-1}).$$