Foundations of Data Science & Machine Learning

Tutotial 08

April 30, 2021

Question 1. Let G be a strongly connected digraph and P be the transition probability matrix of a random walk on G. Show that the stationary distribution π for P is strictly positive, i.e., $\pi_x > 0$ for every vertex $x \in V(G)$.

Solution. Suppose, for the sake of contradiction, there exists a vertex $x \in V(G)$ with $\pi_x = 0$. Since the sum of π is 1, we have at least one vertex $y \in V(G)$ such that $\pi_y > 0$. Since G is strongly connected, there exists a directed path from y to x. Since π-value at the start of the path is positive and that at the end of the path is 0. there exists at least one edge (u, v) in the path with $\pi_u > 0$ and $\pi_v = 0$. This gives the contradiction

$$\begin{aligned} 0 &= \pi_v \\ &= \sum_{z \in N^-(v)} p_{zv}\pi_z \\ &\geq p_{uv}\pi_u \\ &> 0. \end{aligned}$$

Question 2. Let X be a geometric random variable with mean p, that is

$$\forall k \in \mathbb{N}^+,\ P[X = k] = (1-p)^{k-1}p.$$

Show that the expectation $E[X] = 1/p$.

Solution.

$$\begin{aligned} EX &= \sum_{k=1}^{\infty} kP[X = k] && \text{(definition of expectation)} \\ &= p\sum_{k=1}^{\infty} kq^{k-1} && (q = 1-p) \\ EX - qEX &= p\sum_{k=1}^{\infty} q^{k-1} \\ &= p\frac{1}{1-q} \\ &= 1 \\ EX &= \frac{1}{1-q} \\ &= p. \end{aligned}$$

Question 3. Let G be the two-sided *infinite*[1] path whose vertices are labelled by $\mathbb{Z}$ and all edge-weights 1. Consider the simple random walk on G. That is

$$p_{i,j} = \begin{cases} \frac{1}{2}, & j \in \{i-1, i+1\}, \\ 0, & \text{otherwise.} \end{cases}$$

1. Show that this random walk has no stationary distribution. (*Hint.* Think locally.)
2. Show that $h(0,1)$, the hitting time from 0 to 1, is ∞. (*Hint.* Is $h(i, i+1)$ the same for all i?)

Solution. Suppose, for the sake of contradiction, that π is a stationary distribution for this random walk. Then $\forall i \in \mathbb{Z}$,

$$\begin{aligned} \pi_i &= \frac{1}{2}\pi_{i-1} + \frac{1}{2}\pi_{i+1} \\ 2\pi_i &= \pi_{i-1} + \pi_{i+1} \\ \pi_i - \pi_{i-1} &= \pi_{i+1} - \pi_i. \end{aligned}$$

That is, π is an infinite arithmetic progression. If the common difference is not zero, then it will eventually go more than 1 which contradicts the assumption that π is a probability distribution. If the common difference is 0, the $\sum_{i\in\mathbb{Z}} \pi_i$ is either 0 or ∞ and not 1 as needed.

Since every point on the two-sided infinite path looks the same, $h(i, i+1)$ should be same, say h for all $i \in \mathbb{Z}$.

$$\begin{aligned} h(i, i+1) &= \frac{1}{2}1 + \frac{1}{2}h(i-1, i+1) \\ &= \frac{1}{2}1 + \frac{1}{2}(h(i-1, i) + h(i, i+1)) \\ h &= \frac{1}{2} + \frac{1}{2}(h + h) \\ h &= \frac{1}{2} + h, \end{aligned}$$

which cannot happen for any finite h. Hence h is infinite.

Paradox? If the random walk above starts at 0, then we expect it to hit either 1 or -1 in one step. But the expected time to hit 1 is ∞ and the expected time to hit -1 is also (by a similar argument) ∞. How do you get your head around it?

[1] This is an out of syllabus question since I promised we will only study finite graphs. But this gives you a feel of how things start getting different when we make the leap to infinity.