Stochastic Processes

Derek Li

Contents

1 Markov Chain Probabilities

1.1 Markov Chain

Definition 1.1. A discrete-time, discrete-space, time-homogeneous ***Markov chain*** is specified by three ingredients:
(i) A ***state space*** S, any non-empty finite or countable set.
(ii) ***Initial probabilities*** $\{v_i\}_{i\in S}$, where v_i is the probability of starting at i (at time 0). (So $v_i \geqslant 0$ and $\sum\limits_i v_i = 1$.)
(iii) ***Transition probabilities*** $\{p_{ij}\}_{i,j\in S}$, where p_{ij} is the probability of jumping to j if you start at i. (So, $p_{ij} \geqslant 0$ and $\sum\limits_j p_{ij} = 1, \forall i$.)

Note. (1) Given any Markov chain, let X_n be the Markov chain's state at time n and thus $X_0, X_1, \cdots$ are random variables.
(2) At time 0, we have $P(X_0 = i) = v_i, \forall i \in S$.
(3) p_{ij} can be interpreted as conditional probabilities, i.e., if $P(X_n = i) > 0$, then

$$P(X_{n+1} = j|X_n = i) = p_{ij}, \forall i, j \in S, n = 0, 1, \cdots,$$

which does not depend on n because of time-homogeneous property.
(4) The probabilities at time $n+1$ depend only on the state at time n, i.e.,

$$P(X_{n+1} = j|X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n) = P(X_{n+1} = j|X_n = i_n) = p_{i_n j},$$

which is called the Markov property.
(5) The joint probabilities can be computed by relating them to conditional probabilities:

$$\begin{aligned} P(X_0 = i_0, X_1 = i_1, \cdots, X_n = i_n) &= P(X_0 = i_0)P(X_1 = i_1|X_0 = i_0) \cdots P(X_n = i_n|X_{n-1} = i_{n-1}) \\ &= v_{i0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}, \end{aligned}$$

which completely defines the probabilities of the sequence $\{X_n\}_{n=0}^{\infty}$. The random sequence $\{X_n\}_{n=0}^{\infty}$ is the Markov chain.

Example 1.1 (Bernoulli Process)**.** Let $0 < p < 1$. Suppose repeatedly flip a p-coin at times $1, 2, \cdots$. Let X_n be the number of heads on the first n flips, then $\{X_n\}$ is a Markov chain, with $S = \{0, 1, \cdots\}, X_0 = 0$ (i.e., $v_0 = 1$ and $v_i = 0, \forall i \neq 0$), and

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

Example 1.2 (Simple Random Walk)**.** Let $0 < p < 1$. Suppose repeatedly bet \$1. Each time, you have probability p of winning \$1 and probability $1 - p$ of losing \$1. Let X_n be the net gain after n bets, then $\{X_n\}$ is a Markov chain, with $S = \mathbb{Z}, X_0 = a$ for some $a \in \mathbb{Z}$ (i.e., $v_a = 1$), and

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

If $p = \frac{1}{2}$, we call it simple symmetric random walk since $p = 1 - p$.

Example 1.3 (Ehrenfest's Urn)**.** Suppose we have d balls, divided into two urns. At each time, we choose one of the d balls uniformly at random, and move it to the other urn. Let X_n be the number of balls in Urn 1 at time n, then $\{X_n\}$ is a Markov chain, with $S = \{0, 1, \cdots, d\}$, and

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

1.2 Multi-Step Transitions

Let $\mu_i^{(n)} = P(X_n = i)$ be the probabilities at time n : at time 0, $\mu_i^{(0)} = P(X_0 = i) = v_i$; at time 1, $\mu_j^{(1)} = P(X_1 = j) = \sum\limits_{i \in S} P(X_0 = i, X_1 = j) = \sum\limits_{i \in S} v_i p_{ij} = \sum\limits_{i \in S} \mu_i^{(0)} p_{ij}$ by the Law of Total Probability; at time 2, $\mu_k^{(2)} = \sum\limits_{i \in S} \sum\limits_{j \in S} v_i p_{ij} p_{jk}$, etc.

Let $m = |S|$ be the number of elements in S (could be infinity), $v = (v_1, v_2, \cdots, v_m)$ be a $1 \times m$ row vector, $\mu^{(n)} = (\mu_1^{(n)}, \mu_2^{(n)}, \cdots, \mu_m^{(n)})$ be a $1 \times m$ row vector, and

$$P = (p_{ij}) = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}$$

be an $m \times m$ matrix. Therefore, in matrix form: $\mu^{(1)} = vP = \mu^{(0)} P, \mu^{(2)} = vPP = vP^2 = \mu^{(0)} P^2$. By induction, we have

$$\mu^{(n)} = vP^n = \mu^{(0)} P^n, n \in \mathbb{N}.$$

By convention, let $P^0 = I$, then $\mu^{(n)} = vP^n$ holds for $n = 0$.

Another way to track the probabilities of a Markov chain is with n-step transitions

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i).$$

Since the chain is time-homogeneous, $p_{ij}^{(n)} = P(X_{m+n} = j | X_m = i), \forall m \in \mathbb{N}$. Note that $p_{ij}^{(n)} \geqslant 0$ and $\sum\limits_{j \in S} p_{ij}^{(n)} = \sum\limits_{j \in S} P_i(X_n = j) = P_i(X_n \in S) = 1$. We have $p_{ij}^{(1)} = P(X_1 = j | X_0 = i) = p_{ij}$, and $p_{ij}^{(2)} = P(X_2 = j | X_0 = i) = \sum\limits_{k \in S} P(X_2 = j, X_1 = k | X_0 = i) = \sum\limits_{k \in S} p_{ik} p_{kj}, p_{ij}^{(3)} = \sum\limits_{k \in S} \sum\limits_{l \in S} p_{ik} p_{kl} p_{lj}$, etc. Therefore, in matrix form: $P^{(2)} = (p_{ij}^(2)) = PP = P^2, P^{(3)} = P^3$. By induction we have

$$P^{(n)} = P^n, n \in \mathbb{N}.$$

By convention, let $P^{(0)} = I$, then $P^{(n)} = P^n$ holds for $n = 0$.

Theorem 1.1 (Chapman-Kolmogorov Equations)**.**

$$p_{ij}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}, p_{ij}^{m+s+n} = \sum_{k \in S} \sum_{l \in S} p_{ik}^{(m)} p_{kl}^{(s)} p_{lj}^{(n)}, \text{ etc.}$$

Proof. By the Law of Total Probability,

$$p_{ij}^{(m+n)} = P(X_{m+n} = j | X_0 = i) = \sum_{k \in S} P(X_{m+n} = j, X_m = k | X_0 = i) = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}.$$

□

In matrix form: $P^{(m+n)} = P^{(m)}P^{(n)}, P^{(m+s+n)} = P^{(m)}P^{(s)}P^{(n)}$, etc.

Theorem 1.2 (Chapman-Kolmogorov Inequality)**.**

$$p_{ij}^{(m+n)} \geqslant p_{ik}^{(m)} p_{kj}^{(n)},$$

for any fixed state $k \in S$, etc.

1.3 Recurrence and Transience

Let $N(i) = |\{n \geqslant 1 : X_n = i\}|$ be the total number of times that the chain hits i (not counting time 0) and so $N(i)$ is a random variable, possibly infinite. Let f_{ij} be the ***return probability*** from i to j, i.e., f_{ij} is the probability, starting from i, that the chain will eventually visit j at least once:

$$f_{ij} := P_i(X_n = j \text{ for some } n \geqslant 1) = P_i(N(j) \geqslant 1).$$

Thus, we have

$$1 - f_{ij} = P_i(X_n \neq j \text{ for all } n \geqslant 1).$$

Also, we have

$$P_i(\text{Chain will eventually visit } j, \text{ and then eventually visit } k) = f_{ij} f_{jk}, \text{ etc.}$$

Hence, $P_i(N(i) \geqslant k) = (f_{ii})^k, P_i(N(j) \geqslant k) = f_{ij}(f_{jj})^{k-1}$.

Property 1.1. $f_{ik} \geqslant f_{ij} f_{jk}$, etc.

Definition 1.2. A state i of a Markov chain is ***recurrent*** or ***persistent*** if

$$P_i(X_n = i \text{ for some } n \geqslant 1) = 1, \text{ i.e., } f_{ii} = 1.$$

Otherwise, if $f_{ii} < 1$, then i is ***transient***.

Theorem 1.3 (Recurrent State Theorem)**.** State i is recurrent iff $P_i(N(i) = \infty) = 1$ iff $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$. State i is transient iff $P_i(N(i) = \infty) = 0$ iff $\sum_{n=1}^{\infty} p_{ii}^{(n)} < \infty$.