An introduction to Markov chains

Jie Xiong

Department of Mathematics
The University of Tennessee, Knoxville

[NIMBioS, March 16, 2011]

Mathematical biology (WIKIPEDIA)

Markov chains also have many applications in biological modelling, particularly population processes, which are useful in modelling processes that are (at least) analogous to biological populations. The Leslie matrix is one such example, though some of its entries are not probabilities (they may be greater than 1). Another example is the modeling of cell shape in dividing sheets of epithelial cells. Yet another example is the state of Ion channels in cell membranes.
Markov chains are also used in simulations of brain function, such as the simulation of the mammalian neocortex.

Outline

1. Basic definitions
2. State classification
3. Stationary distributions and limit behavior
4. Simple random walk
5. Continuous time
6. Poisson process
7. Branching process

1. Basic definitions

A sequence of random variables $\{X_n: \ n = 0, 1, 2, \cdots\}$

State space $\mathcal{S} = \{1, 2, \cdots, N\}$ or $\mathcal{S} = \mathbb{N}$

Definition

$\{X_n\}$ is a Markov chain if

$$\begin{aligned} & \mathbb{P}\left(X_{n+1} = i_{n+1} | X_0 = i_0, \ \cdots, \ X_n = i_n\right) \\ = \ & \mathbb{P}\left(X_{n+1} = i_{n+1} | X_n = i_n\right). \end{aligned}$$

If $\mathcal{S} = \{1, 2, \cdots, N\}$, $\{X_n\}$ is a finite Markov chain.

Definition (1-step transition)

$$p_{ij}(n) = \mathbb{P}\left(X_{n+1} = j | X_n = i\right), \qquad i,\ j \in \mathcal{S}$$

is the 1-step transition probability at time n from i to j. If it does not depend on n, $\{X_n\}$ is a time-homogeneous Markov chain.

Denote the transition matrix by

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} & \cdots \\ p_{21} & p_{22} & p_{23} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

Note that

$$\sum_{j \in \mathcal{S}} p_{ij} = 1, \qquad i \in \mathcal{S}.$$

Definition (n-step transition)

$$p_{ij}^{(n)} = \mathbb{P}\left(X_n = j | X_0 = i\right), \qquad i,\ j \in \mathcal{S}$$

is the n-step transition probability from i to j.

Denote the n-step transition matrix by $P^{(n)}$.
Note that

$$P^{(0)} = I \text{ and } P^{(1)} = P.$$

Theorem (Chapman-Kolmogorov equation)

$$P^{(m+n)} = P^{(m)} P^{(n)}, \qquad \forall\, m,\; n \in \mathbb{N}.$$

Proof:

$$\begin{aligned}
p_{ij}^{(m+n)} &= \mathbb{P}\left(X_{m+n} = j | X_0 = i\right) \\
&= \sum_{k \in \mathcal{S}} \mathbb{P}\left(X_m = k,\; X_{m+n} = j | X_0 = i\right) \\
&= \sum_{k \in \mathcal{S}} \mathbb{P}\left(X_m = k | X_0 = i\right) \mathbb{P}\left(X_{m+n} = j | X_m = k,\; X_0 = i\right) \\
&= \sum_{k \in \mathcal{S}} p_{ik}^{(m)} \mathbb{P}\left(X_{m+n} = j | X_m = k\right) \\
&= \sum_{k \in \mathcal{S}} p_{ik}^{(m)} p_{kj}^{(n)}.
\end{aligned}$$

Corollary

$$P^{(n)} = P^n$$

Proof: (Induction) If $n = 1$, $P^{(1)} = P^1$.
Assume $P^{(k)} = P^k$.
Let $n = k + 1$. Then

$$P^{(k+1)} = P^{(k)} P^1 = P^k P = P^{k+1}$$

2. State classification

Definition

Let $i,\ j \in \mathcal{S}$. j is accessible from i (denote $i \to j$) if $p_{ij}^{(n)} > 0$ for some n.
i and j communicate (denote $i \leftrightarrow j$) if $i \to j$ and $j \to i$.

Note that "communicate" is an equivalence relation, namely

- reflexivity: $i \leftrightarrow i$ ($p_{ii}^{(0)} = 1 > 0$)
- symmetry: $i \leftrightarrow j$ implies $j \leftrightarrow i$
- transitivity: $i \leftrightarrow j$ and $j \leftrightarrow k$ imply $i \leftrightarrow k$

$\mathcal{S}$ is divided into equivalent classes.

Definition

- Each equivalence class is a class of the MC
- If there is only one class, the MC is irreducible
- Class C is closed if

$$p_{ij} = 0, \qquad \forall\ i \in C \text{ and } j \notin C.$$

Example

$$P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{2}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Then

$$C_1 = \{1,2\}, \quad C_2 = \{3\}, \quad C_3 = \{4,5\}$$

C_1, C_3 are closed. C_2 is not closed.

Definition

Period of i is

$$d(i) = gcd\{n \geq 1 : p_{ii}^{(n)} > 0\}, \qquad gcd\emptyset = 0$$

If $d(i) = 1$, i is aperiodic.

Theorem

If $i \leftrightarrow j$, then $d(i) = d(j)$.

Proof: Suppose $d(i) > 0$, $p_{ij}^{(m)} > 0$ and $p_{ji}^{(n)} > 0$.
If $p_{ii}^{(s)} > 0$, then

$$p_{jj}^{(n+s+m)} \geq p_{ji}^{(n)} p_{ii}^{(s)} p_{ij}^{(m)} > 0$$

Similarly,

$$p_{jj}^{(n+2s+m)} > 0.$$

Thus, $d(j)$ divides $n + s + m$ and $n + 2s + m$. So, $d(j)$ divides s.
Thus, $d(j)|d(i)$. By symmetry, $d(i) = d(j)$.

Definition

Let

$$T_i = \inf\{m \geq 1 : \; X_m = i\}$$

and

$$f_{ii}^{(n)} = \mathbb{P}(T_i = n | X_0 = i).$$

i is recurrent if $\sum_{n\geq 1} f_{ii}^{(n)} = 1$. Otherwise, it is transient.

Definition

Suppose i is recurrent. If

$$\mu_{ii} \equiv \mathbb{E}(T_i) = \sum_{n\geq 1} n f_{ii}^{(n)} < \infty,$$

then i is positive recurrent. Otherwise, it is null recurrent.

For $0 < s < 1$, let

$$F_{ii}(s) = \sum_{n=1}^{\infty} f_{ii}^{(n)} s^n = \mathbb{E} s^{T_i}$$

and

$$P_{ii}(s) = \sum_{n=0}^{\infty} p_{ii}^{(n)} s^n.$$

Lemma

$$P_{ii}(s)(1 - F_{ii}(s)) = 1.$$

Proof: As

$$p_{ii}^{(n)} = \sum_{m=1}^{n} f_{ii}^{(m)} p_{ii}^{(n-m)},$$

we have

$$\begin{aligned} P_{ii}(s) &= 1 + \sum_{n=1}^{\infty} \sum_{m=1}^{n} f_{ii}^{(m)} p_{ii}^{(n-m)} s^n \\ &= 1 + \sum_{m=1}^{\infty} f_{ii}^{(m)} \sum_{k=0}^{\infty} p_{ii}^{(k)} s^k s^m \\ &= 1 + F_{ii}(s) P_{ii}(s). \end{aligned}$$

Theorem

i is recurrent iff

$$\sum_n p_{ii}^{(n)} = \infty.$$

Proof: Note that

$$\sum_n p_{ii}^{(n)} = P_{ii}(1-) = \frac{1}{1 - F_{ii}(1-)}$$

$$= \frac{1}{1 - \sum_n f_{ii}^n}.$$

Corollary

If $i \leftrightarrow j$, then i is recurrent iff j is recurrent.

Key for proof:

$$p_{jj}^{(n+s+m)} \geq p_{ji}^{(n)} p_{ii}^{(s)} p_{ij}^{(m)}.$$

Corollary

Every recurrent class is closed.

Proof: If $i \in C$, $j \notin C$ and $i \to j$, then

$$\mathbb{P}(\exists\, n,\ X_n = i | X_0 = i) \neq 1.$$

This contradicts from i being recurrent.

3. Stationary distributions and limit behavior

Theorem

For an irreducible, aperiodic, recurrent MC, we have

$$\lim_{n\to\infty} p_{ij}^{(n)} = \frac{1}{\mu_{jj}}, \qquad \forall\ i,\ j \in \mathcal{S}.$$

Example Let

$$P = \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \end{bmatrix}$$

Let

$$T = \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix} \text{ and } T^{-1} = \frac{1}{5} \begin{bmatrix} 2 & 3 \\ 1 & -1 \end{bmatrix}$$

Then

$$P = T \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{6} \end{bmatrix} T^{-1}$$

Thus

$$P^n = T \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{6^n} \end{bmatrix} T^{-1} \to T \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} T^{-1} = \frac{1}{5} \begin{bmatrix} 2 & 3 \\ 2 & 3 \end{bmatrix}$$

Hence

$$\mu_{11} = \frac{5}{2} \text{ and } \mu_{22} = \frac{5}{3}$$

Corollary

For an irreducible, d-periodic, recurrent MC, we have

$$\lim_{n\to\infty} p_{ii}^{(nd)} = \frac{1}{\mu_{ii}}, \qquad \forall\ i \in \mathcal{S}.$$

Proof: $X_{nd},\ n = 0, 1, 2, \cdots$ is then an irreducible, aperiodic, recurrent MC

Let $(\pi_i)_{i\in\mathcal{S}}$ be the initial distribution, i.e.,

$$\mathbb{P}(X_0 = i) = \pi_i, \qquad i \in \mathcal{S}.$$

Then,

$$\begin{aligned}\mathbb{P}(X_n = j) &= \sum_{i\in\mathcal{S}} \mathbb{P}(X_n = j | X_0 = i) P(X_0 = i) \\ &= \sum_{i\in\mathcal{S}} \pi_i p_{ij}^{(n)}.\end{aligned}$$

Theorem

The distribution of X_n *is* πP^n.

Definition

(π_i) is a stationary distribution if

$$\pi P = \pi$$

Theorem

For an irreducible, aperiodic, recurrent MC,

$$\pi_i = \frac{1}{\mu_{ii}}, \qquad i \in \mathcal{S}$$

is the unique stationary distribution.

Proof: If (π_i) is a stationary distribution, then

$$\sum_{i \in \mathcal{S}} \pi_i p_{ij}^{(n)} = \pi_j.$$

Letting $n \to \infty$, we get

$$\pi_j = \sum_{i \in \mathcal{S}} \pi_i \frac{1}{\mu_{jj}} = \frac{1}{\mu_{jj}}.$$

On the other hand, we note

$$p_{ij}^{(n+1)} = \sum_{k \in \mathcal{S}} p_{ik}^{(n)} p_{kj}.$$

Taking $n \to \infty$, we get

$$\frac{1}{\mu_{jj}} = \sum_{k \in \mathcal{S}} \frac{1}{\mu_{kk}} p_{kj}.$$

So, $(\frac{1}{\mu_{jj}})_{j \in \mathcal{S}}$ is a stationary distribution.

4. Simple random walk

If $X_n = i$, then $X_{n+1} = i - 1$ or $i + 1$ with equal probabilities. Such a MC is a SRW.

$$P = \begin{bmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \frac{1}{2} & 0 & \frac{1}{2} & \cdots & \\ \cdots & \cdots & \frac{1}{2} & 0 & \frac{1}{2} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

Construction of SRW:
Let $\xi_1, \xi_2, \cdots$ be i.i.d.

$$\mathbb{P}(\xi_1 = 1) = \mathbb{P}(\xi_1 = -1) = \frac{1}{2}$$

Define

$$X_n = \sum_{k=1}^{n} \xi_k.$$

Theorem

$\{X_n\}$ is a SRW.

Proof:

$$X_{n+1} = X_n + \xi_{n+1}$$

5. Continuous time MC

State space $\mathcal{S}$ (same)
Time set $[0, \infty)$
A family of random variables $\{X_t : \ t \geq 0\}$

Definition

$\{X_t\}$ is a Markov chain if $\forall\ 0 = t_0 < t_1 < \cdots < t_{n+1}$,

$$\begin{aligned} & \mathbb{P}\left(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \ \cdots, \ X_{t_n} = i_n\right) \\ = \ & \mathbb{P}\left(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n\right). \end{aligned}$$

Transition probabilities (time-homogeneous case)

$$\mathbb{P}(X_t = j | X_s = i) = p_{ij}(t-s).$$

Denote

$$P(t) = (p_{ij}(t))_{i,j\in\mathcal{S}}$$

Theorem (Chapman-Kolmogorov equation)

$$P(t+s) = P(t)P(s), \qquad \forall\, t,\ s \geq 0.$$

It can be proved that $P(t)$ is differentiable. Denote

$$Q = P'(0)$$

Definition

Q is the infinitesimal generator matrix.

Theorem

$$P'(t) = QP(t) = P(t)Q.$$

Proof:

$$\begin{aligned} P'(t) &= \lim_{h\to 0} h^{-1}(P(t+h) - P(t)) \\ &= \lim_{h\to 0} h^{-1}(P(t)P(h) - P(t)) \\ &= P(t) \lim_{h\to 0} h^{-1}(P(h) - I) \\ &= P(t)Q \end{aligned}$$

Next, we define jump time

$$\tau_0 = 0, \qquad \tau_{n+1} = \inf\{t > \tau_n : \ X_t \neq X_{\tau_n}\}.$$

Define

$$Y_n = X_{\tau_n}, \qquad n = 0, 1, 2, \cdots .$$

Theorem

i) Suppose $q_{ii} \neq 0$. Then, given $X_{\tau_n} = i$, the r.v. $\tau_{n+1} - \tau_n$ has exponential distribution with parameter $-q_{ii}$.

ii) $\{Y_n\}$ is a discrete-time MC with 1-step transition matrix P given by

$$p_{ii} = 0, \qquad p_{ij} = \frac{q_{ij}}{-q_{ii}}, \quad i \neq j.$$

6. Poisson process

Example

X_t is the number of certain items (e.g., birth defect, accident, etc) before time t. Then,

$$X_t \sim Poisson(\lambda t)$$

where λ is the average number of items in a unit time.
X_t is a Poisson process. Time gaps between events are i.i.d. exponential random variables with parameter $\theta = \frac{1}{\lambda}$.

Definition

(X_t) is a Poisson process if

- $X_0 = 0$
- $\forall\ 0 = t_0 < t_1 < \cdots < t_n$, the increments

$$X_{t_i} - X_{t_{i-1}}, \qquad i = 1, 2, \cdots, n$$

  are independent.
- $X_t - X_s \sim Poisson(\lambda(t-s))$.

Theorem

For Poisson process, we have

$$q_{ii} = -\lambda, \qquad q_{i,i+1} = \lambda, \qquad i = 0, 1, 2, \cdots.$$

7. Critical binary branching process

Initially, there are X_0 individuals. Each has an exponential clock with para. γ, i.e.,

$$\mathbb{P}(\eta_k > t) = e^{-\gamma t}, \qquad t > 0, \quad k = 1, 2, \cdots, X_0.$$

When the time is up, that individual will split to 2 or die with equal probability. Let X_t be the number of individuals in the population at time t. Then, (X_t) is a MC.

Suppose $X_0 = i$. Then

$$\tau_1 = \min\{\eta_k : k = 1, 2, \cdots, X_0\}.$$

As

$$\mathbb{P}(\tau_1 > t) = \Pi_{k=1}^{i} \mathbb{P}(\eta_k > t) = e^{-i\gamma t}.$$

Thus

$$q_{ii} = -i\gamma.$$

Further,

$$q_{i,i+1} = q_{i,i-1} = \frac{1}{2} i\gamma.$$