Markov Chain lecture notes
Math331, Fall 2008
Instructor: David Anderson

Markov Chains: lecture 1.

Markov chains are probability models for trials of random experiments of great variety, and their defining characteristic is that they allow us to consider situations where the future evolution of the process of interest depends on where it is at present, but not on how it got there. This contrasts with the independent trials models we have considered in the law of large numbers and the central limit theorem. For independent trial processes the possible outcomes of each trial of the experiment are the same and occur with the same probability. Furthermore, what happens on any trial is not affected by what happens on any other trial. With Markov chain models we can generalize this to the extent that we allow the future to depend on the present. We formulate this notion precisely in the following definition.

Definition 1. Let $\{X_n : n \geq 0\}$ be a sequence of random variables with values in a finite or countably infinite set S. Furthermore, assume

$$(*) \quad P(X_{n+1} = y \mid X_0 = x_0, X_1 = x_1, \ldots, X_{n-1} = x_{n-1}, X_n = x) = P(X_{n+1} = y | X_n = x) = p(x, y)$$

for all $n \geq 0$ and all states $x, y, x_0, x_1, \ldots, x_{n-1}$ in S. Then $\{X_n \mid n \geq 0\}$ is called a <u>Markov chain</u> with <u>state space</u> S and <u>transition matrix</u> $P = [p(x, y)]$. The property $(*)$ is called the Markov property.

• Therefore, Markov chains are processes whose future is dependent only upon the present state. Such processes arise abundantly in the natural, mathematical, and social sciences. **Some examples**: Position of a random walker, number of each species of an ecosystem in a given year,...

The <u>initial distribution</u> for the Markov chain is the sequence of probabilities

$$\pi(x) = P\{X_0 = x\}, \quad x \in S,$$

or, in vector form,

$$\pi = [\pi(x) \mid x \in S].$$

Remarks:

(i) $(*)$ can be interpreted as stating that the conditional distribution of the random future state X_{n+1} depends only on the present state X_n and is independent of the past states $X_0, \ldots, X_{n-1}$.

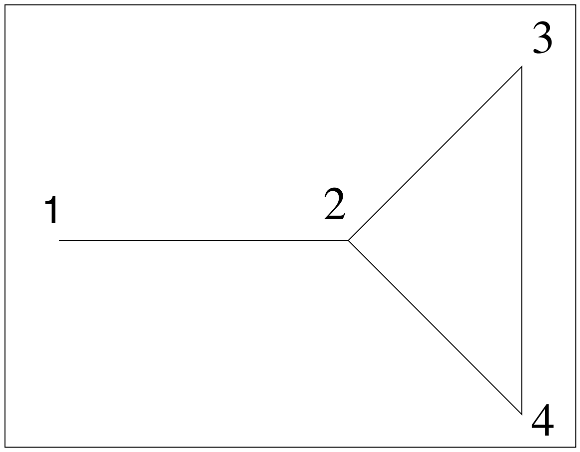
(ii) One can think of a Markov chain as a model for jumping from state to state of S and all jumps are governed by the jump probabilities $p(i, j)$. A frog jumping from one pad to the next in a pond where the pads consist of the states of S and the jumps are taken according to the transition probabilities is a good picture to keep in mind.

(iii) Observe that $p(x, y) \geq 0$ for all $x, y \in S$ and

$$\sum_{y \in S} p(x, y) = 1.$$

(iv) If S has r elements, we will frequently denote S by $\{1, 2, \ldots, r\}$ or $\{0, 1, ..., r-1\}$. This appears in the notation used, sometimes without explicit mention. Hence keep this in mind as you read.

Example 1. A colleague travels between four coffee shops located as follows:



Assume he/she chooses among the paths departing from a shop by treating each path as equally likely. If we model our colleague's journey as a Markov chain, then a suitable state space would be $S = \{1, 2, 3, 4\}$ and the transition matrix is easily seen to be

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

Suppose this person has an initial distribution among the shops as follows:

$$\pi(x) = P(X_0 = x_0) = 1/4, \qquad \text{for } x_0 \in \{1, 2, 3, 4\}$$

or, in vector form,

$$\pi = \left[\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right].$$

Show that

$$P(X_0 = 1, X_1 = 2, X_2 = 4, X_3 = 3) = \pi(1)p(1,2)p(2,4)p(4,3) = 1/24.$$

Using conditioning and the Markov property we see that

$$\begin{aligned} P(X_0 = 1, X_1 = 2, X_2 = 4, X_3 = 3) &= P(X_3 = 3 \mid X_0 = 1, X_1 = 2, X_2 = 4)P(X_0 = 1, X_1 = 2, X_2 = 4) \\ &= p(4,3)P(X_2 = 4 | X_0 = 1, X_1 = 2)P(X_0 = 1, X_1 = 2) \\ &= p(4,3)p(2,4)P(X_1 = 2 | X_0 = 1)P(X_0 = 1) \\ &= p(4,3)p(2,4)p(1,2)\pi(1) \\ &= \pi(1)p(1,2)p(2,4)p(4,3) \\ &= (1/4)(1)(1/3)(1/2) \\ &= 1/24. \end{aligned}$$

Using similar reasoning, we have the general fact that

$$P(X_0 = x_0, X_1 = x_1, \ldots, X_{n-1} = x_{n-1}, X_n = x_n) = \pi(x_0)p(x_0, x_1)p(x_1, x_2) \cdots p(x_{n-1}, x_n).$$

which holds for all $n \geq 0$ and all choices of states in S.

We also have

$$P(X_1 = y) = \sum_{x=1}^{4} P(X_0 = x, X_1 = y) = \sum_{x=1}^{4} \pi(x)p(x,y) = [\pi(1)\ \pi(2)\ \pi(3)\ \pi(4)] \begin{bmatrix} p(1,y) \\ p(2,y) \\ p(3,y) \\ p(4,y) \end{bmatrix},$$

and hence

$$[P(X_1 = 1), P(X_1 = 2), P(X_1 = 3), P(X_1 = 4)] = \pi P.$$

Similarly,

$$\begin{aligned} P(X_2 = 4) &= \sum_{x=1}^{4} P(X_1 = x, X_2 = 4) = \sum_{x=1}^{4} P(X_2 = 4 \mid X_1 = x)P(X_1 = x) \\ &= \sum_{x=1}^{4} P(X_1 = x)p(x,4) \\ &= [P(X_1 = 1), P(X_1 = 2), P(X_1 = 3), P(X_1 = 4)] \begin{bmatrix} p(1,4) \\ p(2,4) \\ p(3,4) \\ p(4,4) \end{bmatrix} = \pi P \begin{bmatrix} p(1,4) \\ p(2,4) \\ p(3,4) \\ p(4,4) \end{bmatrix} \end{aligned}$$

More generally,

$$\begin{aligned} P(X_2 = y) &= \sum_{x=1}^{4} P(X_1 = x)p(x,y) \\ &= [P(X_1 = 1), P(X_1 = 2), P(X_1 = 3), P(X_1 = 4)] \begin{bmatrix} p(1,y) \\ p(2,y) \\ p(3,y) \\ p(4,y) \end{bmatrix} = \pi P \begin{bmatrix} p(1,y) \\ p(2,y) \\ p(3,y) \\ p(4,y) \end{bmatrix} \end{aligned}$$

Thus,

$$[P(X_2 = 1), P(X_2 = 2), P(X_2 = 3), P(X_2 = 4)] = \pi P P = \pi P^2.$$

Continuing in this fashion we get the following general fact.

Theorem 1. *Let P denote the transition matrix of a Markov chain $\{X_n \mid n \geq 0\}$ with initial distribution π. Then*

$$P(X_n = y) = y^{th} \text{ entry of } \pi P^n,$$

and

$$P(X_n = y \mid X_0 = x) = xy^{th} \text{ entry of } P^n.$$

Proof. Assume that $S = \{1, 2, \ldots, r\}$. Let $W_0 = \pi$, $W_1 = [P(X_1 = 1), \cdots, P(X_1 = r)], \ldots, W_n = [P(X_n = 1), \ldots, P(X_n = r)]$. Then,

$$W_1 = \pi P.$$

Similarly,

$$W_2 = \pi P^2, \ldots, W_n = \pi P^n.$$

Hence $P(X_n = y) = y^{th}$ entry of πP^n.

If $\pi = [\pi(1), \ldots, \pi(r)]$ with $\pi(x) = 1$ and $\pi(y) = 0$ if $y \neq x$, then

$$\pi P^n = [p^n(x,1), \ldots, p^n(x,r)]$$

where $p^n(x,y) = P(X_n = y | X_0 = x)$. But note that because of the special form of π, πP^n is simply the x^{th} row of P^n, and so $p^n(x,y)$ is the xy^{th} entry of P^n. □

Exercises:

1. Suppose there are three white and three black balls in two urns distributed so that each urn contains three balls. We say the system is in state i, $i = 0, 1, 2, 3$, if there are i white balls in urn one. At each stage one ball is drawn at random from each urn and interchanged. Let X_n denote the state of the system after the nth draw, and compute the transition matrix for the Markov chain $\{X_n \ : \ n \geq 0\}$.

2. Suppose that whether or not it rains tomorrow depends on previous weather conditions only through whether or not it is raining today. Assume that the probability it will rain tomorrow given it rains today is α and the probability it will rain tomorrow given it is not raining today is β. If the state space is $S = \{0, 1\}$ where state 0 means it rains and state 1 means it does not rain on a given day, find the transition matrix when we model this situation with a Markov chain. If there is a 50-50 chance for rain today, compute the probability it will rain three days from now if $\alpha = 7/10$ and $\beta = 3/10$.