Chapter 8: Markov Chains[1]

Yunghsiang S. Han

Graduate Institute of Communication Engineering,
National Taipei University
Taiwan

E-mail: yshan@mail.ntpu.edu.tw

[1]Modified from the lecture notes by Prof. Mao-Ching Chiu

8.1 Markov Processes

- A random process $X(t)$ is a Markov process if the future of the process given the present is independent of the past. That is, if for arbitrary times $t_1 < t_2 < \cdots < t_k < t_{k+1}$, we have
  - For discrete-valued Markov processes

$$\begin{aligned} & P[X(t_{k+1}) = x_{k+1} | X(t_k) = x_k, \ldots, X(t_1) = x_1] \\ = \; & P[X(t_{k+1}) = x_{k+1} | X(t_k) = x_k]; \end{aligned}$$

  - For continuous-valued Markov process

$$P[a < X(t_{k+1}) \leq b | X(t_k) = x_k, \ldots, X(t_1) = x_1]$$

$$= P[a < X(t_{k+1}) \leq b | X(t_k) = x_k].$$

- The pdf of a Markov process is given by

$$\begin{aligned} & f_{X(t_{k+1})}(x_{k+1} | X(t_k) = x_k, \ldots, X(t_1) = x_1) \\ = & f_{X(t_{k+1})}(x_{k+1} | X(t_k) = x_k). \end{aligned}$$

Example: Consider the sum process:

$$S_n = X_1 + X_2 + \cdots + X_n = S_{n-1} + X_n,$$

where the X_i's are an iid sequence. S_n is a Markov process since

$$\begin{aligned} P[S_{n+1} = s_{n+1} | S_n = s_n, \ldots, S_1 = s_1] &= P[X_{n+1} = s_{n+1} - s_n] \\ &= P[S_{n+1} = s_{n+1} | S_n = s_n]. \end{aligned}$$

Example: Consider the moving average of a Bernoulli sequence:

$$Y_n = \frac{1}{2}(X_n + X_{n-1}),$$

where X_i are independent Bernoulli sequence with $p = 1/2$. We show that Y_n is not a Markov process. The pmf of Y_n is

$$P[Y_n = 0] = P[X_n = 0, X_{n-1} = 0] = 1/4,$$

$$\begin{aligned} P[Y_n = 1/2] &= P[X_n = 0, X_{n-1} = 1] + P[X_n = 1, X_{n-1} = 0] \\ &= 1/2 \end{aligned}$$

and

$$P[Y_n = 1] = P[X_n = 1, X_{n-1} = 1] = 1/4.$$

Now consider

$$
\begin{aligned}
P[Y_n = 1 | Y_{n-1} = 1/2] &= \frac{P[Y_n = 1, Y_{n-1} = 1/2]}{P[Y_{n-1} = 1/2]} \\
&= \frac{P[X_n = 1, X_{n-1} = 1, X_{n-2} = 0]}{1/2} \\
&= \frac{(1/2)^3}{1/2} = 1/4.
\end{aligned}
$$

Suppose that we have additional knowledge about past, then

$$
\begin{aligned}
& P[Y_n = 1 | Y_{n-1} = 1/2, Y_{n-2} = 1] \\
= & \frac{P[Y_n = 1, Y_{n-1} = 1/2, Y_{n-2} = 1]}{P[Y_{n-1} = 1/2, Y_{n-2} = 1]} = 0.
\end{aligned}
$$

Thus

$$P[Y_n = 1 | Y_{n-1} = 1/2] \neq P[Y_n = 1 | Y_{n-1} = 1/2, Y_{n-2} = 1].$$

- A integer-valued Markov random process is called a Markov chain.
- If $X(t)$ is a Markov chain for $t_3 > t_2 > t_1$, then we have

$$P[X(t_3) = x_3, X(t_2) = x_2, X(t_1) = x_1]$$
$$= P[X(t_3) = x_3|X(t_2) = x_2]P[X(t_2) = x_2|X(t_1) = x_1]P[X(t_1) = x_1].$$

- In general,

$$\begin{aligned} & P[X(t_{k+1}) = x_{k+1}, X(t_k) = x_k, \ldots, X(t_1) = x_1] \\ = \; & P[X(t_{k+1}) = x_{k+1}|X(t_k) = x_k]P[X(t_k) = x_k|X(t_{k-1}) = x_{k-1}]\cdots \\ & \times P[X(t_2) = x_2|X(t_1) = x_1]P[X(t_1) = x_1]. \end{aligned}$$

8.2 DISCRETE-TIME MARKOV CHAIN

- Let X_n be a discrete-time integer-valued Markov chain that starts at $n = 0$ with pmf

$$p_j(0) = P[X_0 = j], \qquad j = 0, 1, 2, \ldots .$$

- The joint pmf of the first $n + 1$ values is

$$\begin{aligned} & P[X_n = i_n, \ldots, X_0 = i_0] \\ = \quad & P[X_n = i_n | X_{n-1} = i_{n-1}] \cdots P[X_1 = i_1 | X_0 = i_0] P[X_0 = i_0]. \end{aligned}$$

- Assume that the one-step state transition probabilities are fixed and do not change with time (homogeneous transition probability), that is,

$$P[X_{n+1} = j | X_n = i] = p_{ij} \quad \text{for all } n.$$

- The joint pmf for $X_n, X_{n-1}, \ldots, X_0$ is then given by

$$P[X_n = i_n, \ldots, X_0 = i_0] = p_{i_{n-1}, i_n} \cdots p_{i_0, i_1} p_{i_0}(0).$$

- X_n is completely specified by the initial pmf $p_i(0)$ and the matrix of one-step transition probabilities P:

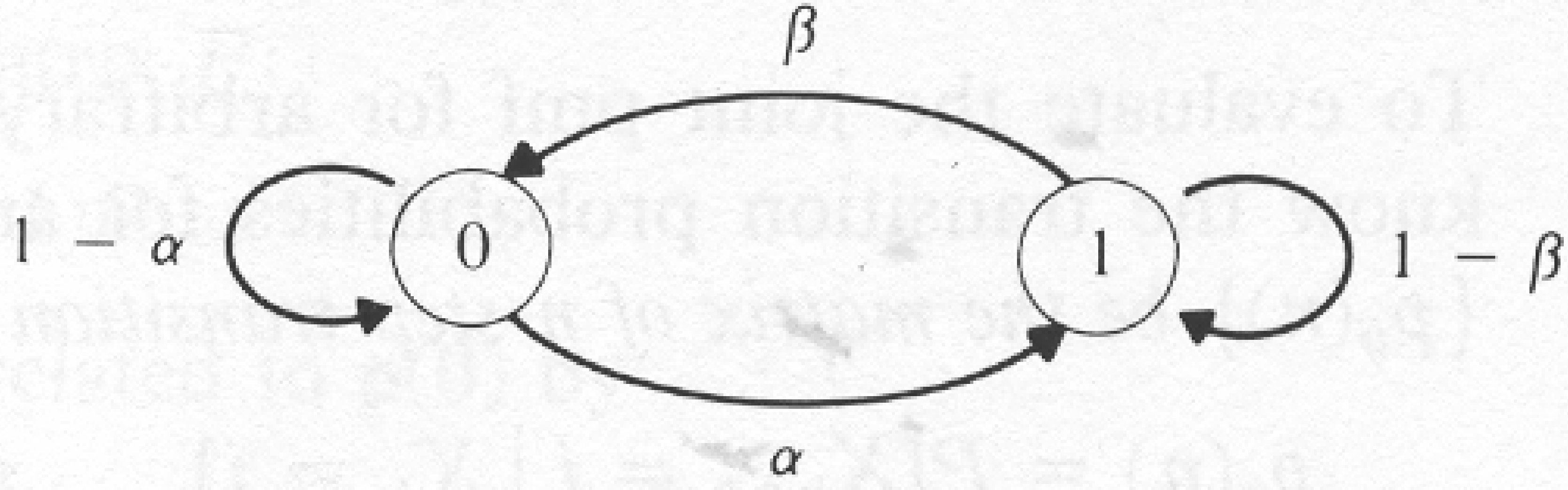
$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ p_{i0} & p_{i1} & p_{i2} & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

- P is called transition probability matrix.
- Each row of P must add to one since

$$1 = \sum_j P[X_{n+1} = j | X_n = i] = \sum_j p_{ij}.$$

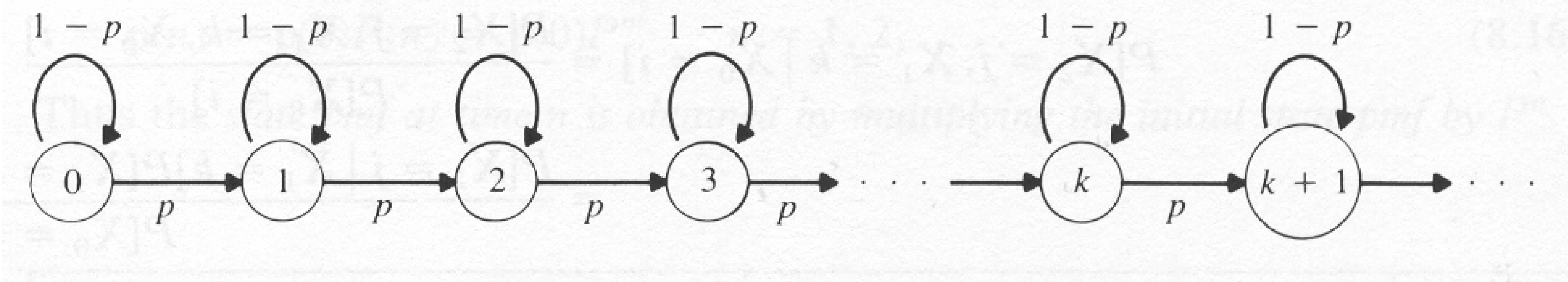
Example: A Markov model for speech:

- Two states: silence and speech activity



Example: Let S_n be the binomial counting process. In one step, S_n can either stay the same or increase by one. The transition probability can be given by

$$P = \begin{bmatrix} 1-p & p & 0 & 0 & \cdots \\ 0 & 1-p & p & 0 & \cdots \\ 0 & 0 & 1-p & p & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix}.$$



The n-step transition probabilities

- Let $P(n) = \{p_{ij}(n)\}$ be the matrix of n-step transition probabilities, where

$$p_{ij}(n) = P[X_{n+k} = j | X_k = i] \quad n \geq 0, \; i, j \geq 0.$$

- Since transition probabilities do not depend on time, we have

$$P[X_{n+k} = j | X_k = i] = P[X_n = j | X_0 = i].$$

- Consider the two-step transition probabilities:

$$P[X_2 = j, X_1 = k | X_0 = i] = \frac{P[X_2 = j, X_1 = k, X_0 = i]}{P[X_0 = i]}$$

$$
\begin{aligned}
&= \frac{P[X_2 = j | X_1 = k] P[X_1 = k | X_0 = i] P[X_0 = i]}{P[X_0 = i]} \\
&= P[X_2 = j | X_1 = k] P[X_1 = k | X_0 = i] \\
&= p_{ik}(1) p_{kj}(1).
\end{aligned}
$$

- 2-step transition probabilities are given by

$$
\begin{aligned}
p_{ij}(2) &= P[X_2 = j | X_0 = i] \\
&= \sum_k P[X_2 = j, X_1 = k | X_0 = i] \\
&= \sum_k p_{ik}(1) p_{kj}(1),
\end{aligned}
$$

- Therefore,

$$P(2) = P(1)P(1) = P^2.$$

- In general, we have

$$P(n) = P^n.$$

State Probabilities

- Let $\boldsymbol{p}(n)$ denote the row vector of state probabilities at time n. The probability $p_j(n)$ is related to $\boldsymbol{p}(n-1)$ by

$$\begin{aligned} p_j(n) &= \sum_i P[X_n = j | X_{n-1} = i] P[X_{n-1} = i] \\ &= \sum_i p_{ij} p_i(n-1). \end{aligned}$$

- In matrix notation we have

$$\boldsymbol{p}(n) = \boldsymbol{p}(n-1) P.$$

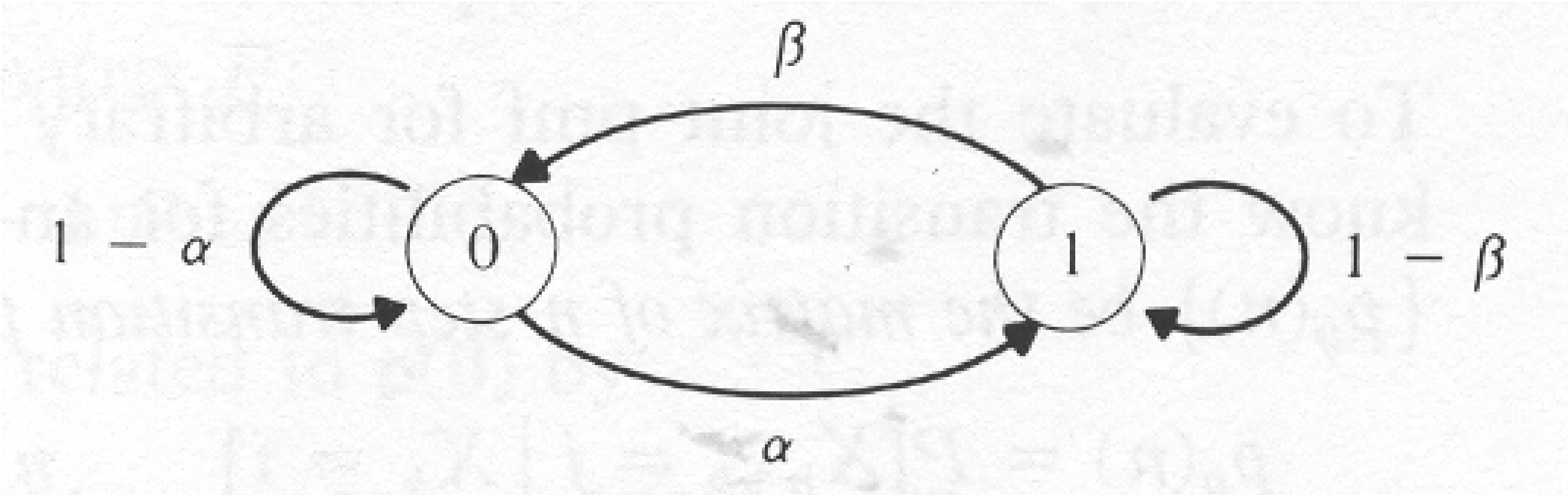
- $p_j(n)$ is related to $\boldsymbol{p}(0)$ by

$$\begin{aligned} p_j(n) &= \sum_i P[X_n = j | X_0 = i] P[X_0 = i] \\ &= \sum_i p_{ij}(n) p_i(0). \end{aligned}$$

- In matrix notation we have

$$\boldsymbol{p}(n) = \boldsymbol{p}(0) P(n) = \boldsymbol{p}(0) P^n.$$

Example: Let $\alpha = 1/10$ and $\beta = 1/5$ for the following Markov chain:



Find $P(n)$ for $n = 2$ and 4.

Sol:

$$P^2 = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}^2 = \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}$$

and

$$P^4 = \begin{bmatrix} 0.83 & 0.17 \\ 0.34 & 0.66 \end{bmatrix}^2 = \begin{bmatrix} 0.7467 & 0.2533 \\ 0.5066 & 0.4934 \end{bmatrix}.$$

Steady State Probabilities

- As $n \to \infty$, the n-step transition probability matrix approaches a matrix in which all the rows are equal to the same pmf

$$p_{ij}(n) \to \pi_j \quad \text{for all } i.$$

- As $n \to \infty$

$$p_j(n) = \sum_i p_{ij} p_i(0) \to \sum_i \pi_j p_i(0) = \pi_j.$$

- As n becomes large, the probability of state j approaches a constant independent of time and the initial state probabilities (equilibrium or steady state).

- Let the pmf $\boldsymbol{\pi} = \{\pi_j\}$. By noting that as $n \to \infty$, $p_j(n) \to \pi_j$ and $p_i(n-1) \to \pi_i$, we have

$$\pi_j = \sum_i p_{ij} \pi_i,$$

which in matrix notation is

$$\boldsymbol{\pi} = \boldsymbol{\pi} P \qquad (n-1 \text{ linearly independent equations}).$$

- The additional equation needed is provided by

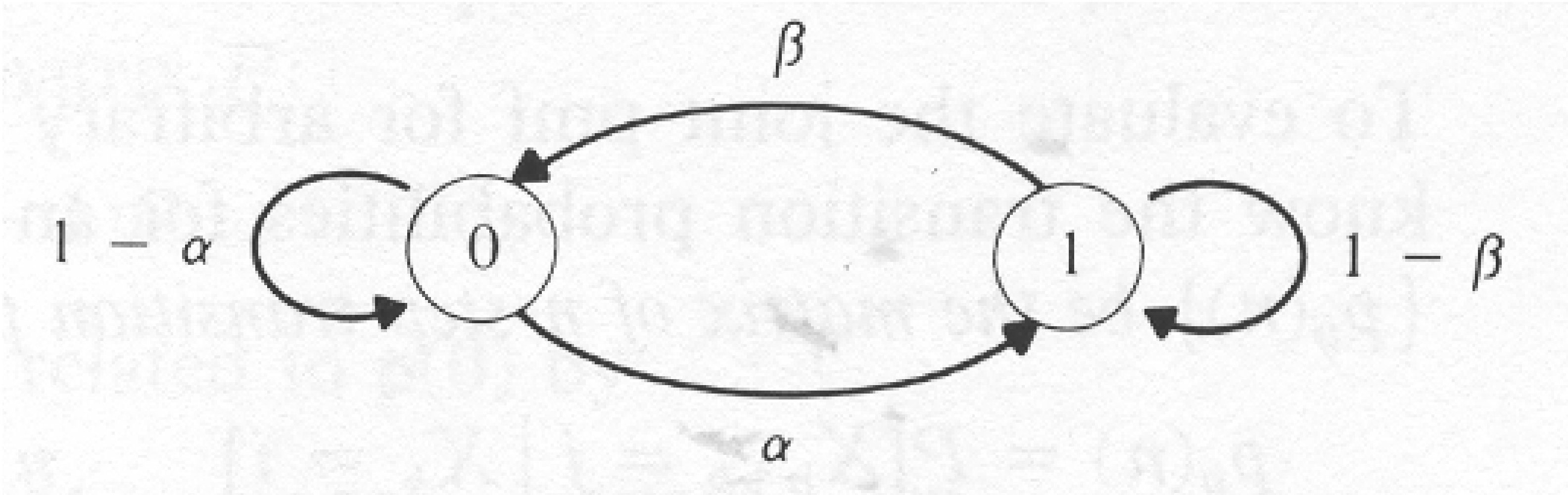
$$\sum_i \pi_i = 1.$$

- $\boldsymbol{\pi}$ is called the stationary state pmf of the Markov chain.

- If we start with $\boldsymbol{p}(0) = \boldsymbol{\pi}$, then

$$\boldsymbol{p}(n) = \boldsymbol{\pi} P^n = \boldsymbol{\pi} \quad \text{– a stationary process.}$$

Example: Find the stationary state pmf for the following Markov chain:



Sol: we have

$$\begin{aligned} \pi_0 &= (1-\alpha)\pi_0 + \beta\pi_i \\ \pi_1 &= \alpha\pi_0 + (1-\beta)\pi_1. \end{aligned}$$

Since $\pi_0 + \pi_1 = 1$,

$$\alpha\pi_0 = \beta\pi_1 = \beta(1 - \pi_0).$$

Thus.

$$\pi_0 = \frac{\beta}{\alpha + \beta}, \qquad \pi_1 = \frac{\alpha}{\alpha + \beta}.$$