Part 2g: Introduction to Markov chains

<u>Textbook</u>: pp. 50-51, 56-57

Introduction to Markov chains

- Markov chains describe processes in (discrete) time.
- The value of the process (the state) varies over time.
- The characteristic feature of Markov processes is that the distribution of the state at time j, for any j, depends on past evolution of the process only via the the value of the process at the previous time $j-1$.

Definition of a Markov chain

Definition 2.13 A stochastic process $X = (X_j)_{j \in \mathbb{N}_0}$ with values in a set S is a *Markov chain*, if

$$P\left(X_j \in A_j \,\middle|\, X_{j-1} \in A_{j-1}, X_{j-2} \in A_{j-2}, \ldots, X_0 \in A_0\right) = P\left(X_j \in A_j \,\middle|\, X_{j-1} \in A_{j-1}\right) \quad (2.2)$$

for all $A_0, A_1, \ldots, A_j \subseteq S$ and all $j \in \mathbb{N}$, that is if the distribution of X_j depends on $X_0, \ldots, X_{j-2}$ only through X_{j-1}. The set S is called the *state space* of X. The distribution of X_0 is called the *initial distribution* of X.

$$f_{X_1, X_2, \ldots, X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdot f_{X_2|X_1}(x_2|x_1) \cdot f_{X_3|X_2,X_1}(x_3|x_2,x_1) \cdot \ldots \times f_{X_n|X_{n-1},\ldots,X_1}(x_n|x_{n-1},\ldots,x_1)$$

$$\underset{\text{Markov}}{=} f_{X_1}(x_1) \cdot f_{X_2|X_1}(x_2|x_1) \cdot f_{X_3|X_2}(x_3|x_2) \cdots \times f_{X_n|X_{n-1}}(x_n|x_{n-1})$$

$f(\cdot)$, $f(\cdot|\cdot) \Leftarrow f(\cdot,\cdot)$

Independence $\Rightarrow$ marginal, Markov $\Rightarrow$ Joint of pairs

Definition 2.16 If the transition probabilities given by the right-hand side of (2.2) do not depend on the time j, the Markov chain X is called *time-homogeneous*.

Discrete state space

- The state space is discrete.
- The state space can be finite, for example $S = \{1,2,3\}$, or it can be countable, for example $S = \{\dots, -1,0,1, \dots\}$.
- Transition distributions can be specified using the transition matrix (possibly infinite):

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

where

$$p_{xy} = P\left(X_j = y \middle| X_{j-1} = x\right)$$

An example

Example 2.17 Consider the state space $S = \{1, 2, 3\}$ and let X be the Markov chain with $X_0 = 1$ where transitions between states have the probabilities:

$$P = \begin{array}{c} 1 \\ 2 \\ 3 \end{array} \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix}.$$
$$\begin{array}{ccc} 1 & 2 & 3 \end{array}$$

$X_0 = 2$

$P(X_1=2, X_2=3, X_3=1, X_4=1, X_5=2 \mid X_0=2)$

$= P_{22} \cdot P_{23} \cdot P_{31} \cdot P_{11} \cdot P_{12} = \frac{1}{2} \frac{1}{2} \cdot 1 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2^4}$

Continuous state space

- The state space is continuous.
- The state space can be an interval or a similar set. It can be one-dimensional or multi-dimensional.
- Frequently, the transition distributions can be specified using transition densities (= kernels):

Definition 2.28 A *transition density* is a map $p: \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ such that:

(a) $p(x, y) \geq 0$ for all $x, y \in \mathbb{R}^d$; and

(b) $\int_{\mathbb{R}^d} p(x, y)\, dy = 1$ for all $x \in \mathbb{R}^d$.

If the Markov chain X can be described by a transition density, then we have

$$P(X_j \in A | X_{j-1} = x) = \int_A p(x, y)\, dy$$

for all $x \in \mathbb{R}^d$.

An example

Example 2.29 On $S = \mathbb{R}$, let $X_0 = 0$ and

$$X_j = \frac{1}{2} X_{j-1} + \varepsilon_j$$

for all $j \in \mathbb{N}$, where $\varepsilon_j \sim \mathcal{N}(0, 1)$ i.i.d. is a Markov chain with state space $S = \mathbb{R}$.

AR(1), $a_1 = \frac{1}{2}$, $\sigma^2 = 1$

$$f(x_1, x_2, \ldots, x_5 \mid X_0 = 0)$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{x_1^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_2 - x_1/2)^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_3 - x_2/2)^2}{2}}$$

$$\times \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_4 - x_3/2)^2}{2}} \times \frac{1}{\sqrt{2\pi}} e^{-\frac{(x_5 - x_4/2)^2}{2}}$$