Markov Chains of Length 3

Random variables form a Markov chain if the distribution of each random variable depends only on the outcome of the random variable that directly precedes it.

Definition: Markov chain (of length 3)

The random variables X, Y, and Z form a Markov chain (notation: $X \to Y \to Z$) if and only if

$$P_{Z|XY} = P_{Z|Y}.$$

In this chapter, we regularly use the shortcut $P_{Z|XY} = P_{Z|Y}$ to denote that this equality should hold for all possible input values, i.e.

$$P_{Z|XY} = P_{Z|Y} \qquad \Leftrightarrow \qquad \forall x \in \mathcal{X}, y \in \mathcal{Y}, z \in \mathcal{Z} : P_{Z|XY}(z|x,y) = P_{Z|Y}(z|y).$$

Proposition

For all random variables X, Y, Z, the following statements are equivalent:

a. $X \to Y \to Z$,
b. $Z \to Y \to X$,
c. $P_{XZ|Y} = P_{X|Y} \cdot P_{Z|Y}$ (that is, $I(X;Z|Y) = 0$).

Proof

We prove that (a) $\Rightarrow$ (b), that (b) $\Rightarrow$ (c), and that (c) $\Rightarrow$ (a). All other directions follow from (combinations of) these three implications.

- [(a) $\Rightarrow$ (b):] Suppose $X \to Y \to Z$. Then

$$
\begin{aligned}
P_{X|YZ} &= \frac{P_{XYZ}}{P_{YZ}} = \frac{P_{XY} \cdot P_{Z|XY}}{P_{YZ}} = \frac{P_{XY} \cdot P_{Z|XY}}{P_Y \cdot P_{Z|Y}} \\
&= \frac{P_{XY} \cdot P_{Z|Y}}{P_Y \cdot P_{Z|Y}} = \frac{P_{XY}}{P_Y} = P_{X|Y},
\end{aligned}
$$

  where we go from the top to the bottom line using the assumption that $X \to Y \to Z$.
- [(b) $\Rightarrow$ (c):] Suppose $Z \to Y \to X$. Then

$$P_{XZ|Y} = \frac{P_{XYZ}}{P_Y} = \frac{P_{X|YZ} \cdot P_{YZ}}{P_Y}$$

$$= \frac{P_{X|Y} \cdot P_{YZ}}{P_Y} = P_{X|Y} \cdot \frac{P_{YZ}}{P_Y} = P_{X|Y} \cdot P_{Z|Y},$$

where we go from the top to the bottom line using the assumption that $Z \to Y \to X$.

- [(c) $\Rightarrow$ (a):] For the final implication, suppose that $P_{XZ|Y} = P_{X|Y} \cdot P_{Z|Y}$. Then

$$P_{Z|YX} = \frac{P_{XYZ}}{P_{XY}} = \frac{P_{XZ|Y} \cdot P_Y}{P_{XY}}$$

$$= \frac{P_{X|Y} \cdot P_{Z|Y} \cdot P_Y}{P_{XY}} = \frac{P_{XY} \cdot P_{Z|Y}}{P_{XY}} = P_{Z|Y},$$

and therefore $X \to Y \to Z$.

Because of the equivalence of items (a) and (b), we often write $X \leftrightarrow Y \leftrightarrow Z$ to denote a Markov chain.

Example: Whispering game

Alice, Bob and Charlie play a game: Alice comes up with a message, and whispers it into Bob's ear. Bob then proceeds to whisper the message into Charlie's ear, who says it out loud. Of course, Bob and Charlie may not hear the message correctly: there is some noise in the communication. Let A be the random variable that contains the message that Alice picks. For simplicity, let's say she picks a uniformly random bit. Let B be the message that Bob heard: it is also a bit, but with probability 0.1, it is not the same as A. More formally,

$$P_{B|A}(0|0) = P_{B|A}(1|1) = 0.9,$$
$$P_{B|A}(1|0) = P_{B|A}(0|1) = 0.1.$$

Finally, let C be the message that Charlie heard: again, let us suppose that it is unequal to the whispered message B with probability 0.1. For the conditional distribution of C, we see that $P_{C|AB}(0|00) = 0.9 = P_{C|B}(0|0)$, and similarly for all other possible values of A, B, and C. Thus, these random variables form a Markov chain: $A \to B \to C$. The proposition above tells us that we can also look at this game in the converse direction: if we are curious about Alice's original message A, we only have to look at Bob's interpretation B of the message. Additionally knowing Charlie's interpretation C does not change the distribution of A.