Graphical Models (PRML)

박태준

March 3, 2023

- We shall find it highly advantageous to augment the analysis using diagrammatic representations of probability distributions, called **probabilistic graphical models.**
- These offer several useful properties:
    1. They provide a simple way to visualize the structure of a probabilistic model and can be used to design and motivate new models.
    2. Insights into the properties of the model, including conditional independence properties, can be obtained by inspection of the graph.
    3. Complex computations, required to perform inference and learning in sophisticated models, can be expressed in terms of graphical manipulations, in which underlying mathematical expressions are carried along implicitly

- A graph comprises **nodes** (also called vertices) connected by **links** (also known as edges or arcs).
- In a probabilistic graphical model,
    - ▷ Each node represents a random variable (or group of random variables)
    - ▷ The links express probabilistic relationships between these variables.
- The graph then captures the way in which the joint distribution over all of the random variables can be decomposed into a product of factors each depending only on a subset of the variables.

- We shall begin by discussing **Bayesian Networks**, also known as directed graphical models, in which the links of the graphs have a particular directionality indicated by arrows. (방향 o)
  - ▷ Directed graphs are useful for expressing causal relationships between random variables
- The other major class of graphical models are **Markov Random Fields**, also known as undirected graphical models, in which the links do not carry arrows and have no directional significance. (방향 x)
  - ▷ Undirected graphs are better suited to expressing soft constraints between random variables.
- For the purposes of solving inference problems, it is often convenient to convert both directed and undirected graphs into a different representation called a **factor graph**.

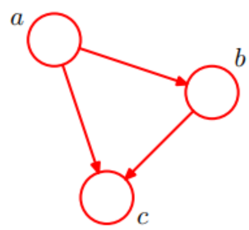
1 Bayesian Networks

- In order to motivate the use of directed graphs to describe probability distributions, consider first an arbitrary joint distribution $p(a, b, c)$ over three variables a, b, and c.
  - ▷ Note that at this stage, we do not need to specify anything further about these variables, such as whether they are discrete or continuous.
- Indeed, one of the powerful aspects of graphical models is that a specific graph can make probabilistic statements for a broad class of distributions.
- We can write the joint distribution in the form

$$\begin{aligned} p(a, b, c) &= p(c|a, b)p(a, b) \\ &= p(c|a, b)p(b|a)p(a) \end{aligned} \tag{8.2}$$

  - ▷ Note that this decomposition holds for any choice of the joint distribution.

Figure 8.1 A directed graphical model representing the joint probability distribution over three variables a, b, and c, corresponding to the decomposition on the right-hand side of (8.2).



- We introduce a node for each of the random variables a, b, and c and associate each node with the corresponding conditional distribution.
- Then, for each conditional distribution we add directed links (arrows) to the graph from the nodes corresponding to the variables on which the distribution is conditioned.
- Thus, for the factor $p(c|a, b)$, there will be links from nodes a and b to node c, whereas for the factor $p(a)$ there will be no incoming links.
- If there is a link going from a node a to a node b, then we say that node a is the **parent** of node b, and we say that node b is the **child** of node a.

- Indeed, in making the decomposition in (8.2), we have implicitly chosen a particular ordering, namely a, b, c, and had we chosen a different ordering we would have obtained a different decomposition and hence a different graphical representation.

- For the moment let us extend the example of Figure 8.1 by considering the joint distribution over K varibles given by $p(x_1, \cdots, x_K)$.