PRML Note C08 Graphical Models

Yang Zhao

Department of Automation, Tsinghua University

- The properties of probabilistic graphical models
 - 1. 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 spphisticated models, can be expressed in terms of graphical manipulations, in which underlying mathematical expressions are carried along implicitly.
- The graph comprises vertices connected by edges, where the vertices represents a random variable and the edges express probabilistic relationships between the variables.
- The categories of the graphical models
 - 1. Bayesian networks: directed graphical models.
 - 2. Markov random fields: undirected graphical models.
 - 3. factor graph: be used to do the inference.
- Graph terminology: a graph G = (V, E) can be represented by adjacency matrix, in which we write $\mathbf{L}(s,t) = 1$ if $(s,t) \in E$. We say the graph is undirected if $\mathbf{L} = \mathbf{L}^T$, otherwise it is directed. We usually assume $\mathbf{L}(s,s) = 0$ which means there are no self loops.
 - 1. Parent: For a directed graph, the parents of a node is the set of all nodes that feed into it: $Pa(s) \triangleq \{t : \mathbf{L}(t,s) = 1\}.$
 - 2. Child: For a directed graph, the children of a node is the set of all nodes that feed out of it: $Ch(s) \triangleq \{t : \mathbf{L}(s,t) = 1\}.$
 - 3. Family: For a directed graph, the family of a node is the node and its parents, $Fam(s) = \{s\} \cup Pas$.
 - 4. Root: For a directed graph, a root is a node with no parents.
 - 5. Leaf: For a directed graph, a leaf is a node with no children.

- 6. Ancestors: For a directed graph, the ancestors are the parents, grand- parents, etc of a node.
- 7. Descendants: For a directed graph, the descendants are the children, grand-children, etc of a node.
- 8. *Neighbors*: For any graph, we define the neighbors of a node as the set of all immediately connected nodes.
- 9. Degree: The degree of a node is the number of neighbors. For directed graphs, we speak of the in-degree and out-degree, which count the number of parents and children.
- 10. Cycle or loop: For any graph, we define a cycle or loop to be a series of nodes such that we can get back to where we started.
- 11. DAG: A directed acyclic graph is a directed graph with no directed cycles.
- 12. Topological ordering: For a DAG, a topological ordering or total ordering is a numbering of the noders such that parents have lower numbers than their children.
- 13. Path or trail: A path or trail $s \rightsquigarrow t$ is a series of directed edges leading from s to t.
- 14. Tree: An undirected tree is an undirected graph with no cycles. A directed tree is a DAG in which there are no directed cycles. If we allow a node to have multiple parents, we call it a polytree.
- 15. Forest: A set of trees.
- 16. Subgraph: A subgraph G_A is the graph created by using the nodes in A and their corresponding edges, $G_A = (V_A, E_A)$.
- 17. Clique: For an undirected graph, a clique is a set of nodes that are all neighbors of each other.

1 Bayesian Networks

• For a graph with K nodes, the joint distribution is given by

$$p(\boldsymbol{x}) = \prod_{k=1}^{K} p(x_k | Pa(x_k))$$

This key equation expresses the *factorization* properties of the joint distribution for a directed graphical model.

2 Conditional Independence

3 Markov Random Fields

4 Inference in Graphical Models

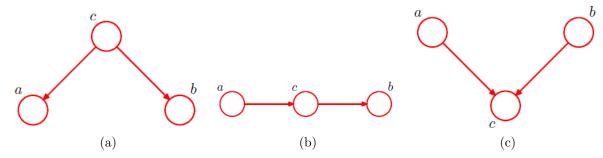


Figure 1: Three cases of 3-node graphs. For the figure 1a, the node c is said to be tail-to-tail with respect to the path from node a to b. For the figure 1b, the node c is said to be head-to-tail with respect to the path from node a to b. For the figure 1c, the node c is said to be head-to-head with respect to the path from node a to b.