Conclusion of chapter 3: Belief Network

yiqing.tao

October 20, 2016

1 summary

1. We can reason with certain or uncertain evidence using repeated application of Bayes rule.
2. A belief network represents a factorisation of a distribution into conditional probabilities of variables de- pendent on parental variables.
3. Belief networks correspond to directed acyclic graphs.
4. Variables are conditionally independent $x \Pi y | z$, if $p(x, y|z) = p(x|z)p(y|z)$; the absence of a link in a belief network corresponds to a conditional independence statement.
5. If in the graph representing the belief network, two variables are independent, then they are independent in any distribution consistent with the belief network structure.
6. Belief networks are natural for representing causal influences.
7. Causal questions must be addressed by an appropriate causal model.

2 belief network(BN)

1. Definition 3.1 (Belief network). A belief network is a distribution of the form
$$p(x_1, \ldots, x_D) = \prod_{i=1}^{n} p(x_i|pa(x_i)) \quad (1)$$
   - where $pa(x_i)$represent the parental variables of variable x_i. Represented as a directed graph, with an arrow pointing from a parent variable to child variable, a belief network corresponds to a Directed Acyclic Graph (DAG), with the i^{th} node in the graph corresponding to the factor $p(x_i|pa(x_i))$.

2.
$$p(x_1, \ldots, x_n) = p(x_n) \prod_{i=1}^{n-1} p(x_i | x_{i+1}, \ldots, x_n) \tag{2}$$

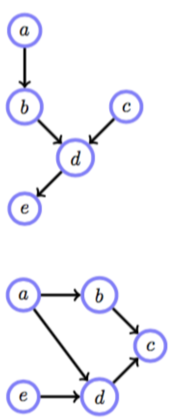
- The observation that any distribution may be written in the cascade form.Gives an algorithm for constructing a BN on variables $x_1, \ldots, x_n$: write down the n node cascade graph; label the nodes with the variables in any order; now each successive independence statement corresponds to deleting one of the edges. More formally, this corresponds to an ordering of the variables which, without loss of generality, we may write as $x_1, \ldots, x_n$.

3. The representation of any BN is therefore a Directed Acyclic Graph (DAG).

4. Conditional independence & Graphical Dependence

   - Remark 3.4 (Graphical Dependence). Belief network (graphs) are good for encoding conditional indepen- dence but are not well suited for encoding dependence. For example, consider the graph a ! b. This may appear to encode the relation that a and b are dependent. However, a specific numerical instance of a belief network distribution could be such that $p(b|a) = p(b)$, for which $a\,\mathrm{II}\,b$. The lesson is that even when the DAG appears to show ?graphical? dependence, there can be instances of the distributions for which depen- dence does not follow. The same caveat holds for Markov networks, section(4.2). We discuss this issue in more depth in section(3.3.5).

5. Definition 3.2 (collide). Given a path P, a collider is a node c on P with neighbours a and b on P such that $a \to c \leftarrow b$. Note that a collider is path specific, see fig(3.8).



(a): The variable d is a collider along the path $a - b - d - c$, but not along the path $a - b - d - e$. Is $a \perp\!\!\!\perp e \,|\, b$? a and e are *not* d-connected since there are no colliders on the only path between a and e, and since there is a non-collider b which is in the conditioning set. Hence a and e are d-separated by b, $\Rightarrow a \perp\!\!\!\perp e | b$.

(b): The variable d is a collider along the path $a - d - e$, but not along the path $a - b - c - d - e$. Is $a \perp\!\!\!\perp e | c$? There are two paths between a and e, namely $a - d - e$ and $a - b - c - d - e$. The path $a - d - e$ is not blocked since although d is a collider on this path and d is not in the conditioning set, we have a descendant of the collider d in the conditioning set, namely c. For the path $a - b - c - d - e$, the node c is a collider on this path and c is in the conditioning set. For this path d is not a collider. Hence this path is not blocked and a and e are (graphically) dependent given c.

Figure 3.8: Collider examples for d-separation and d-connection.

6. Definition 3.4 (d-connection, d-separation). If G is a directed graph in which X, Y and Z are disjoint sets of vertices, then X and Y are d-connected by Z in G if and only if there exists an undirected path U between some vertex in X and some vertex in Y such that for every collider C on U, either C or a descendent of C is in Z, and no non-collider on U is in Z. X and Y are d-separated by Z in G if and only if they are not d-connected by Z in G. One may also phrase this as follows. For every variable $x \in X$ and $y \in Y$, check every path U between x andy. A path U is said to be blocked if there is a node ω on U such that either:

   1. ω is a collider and neither ω nor any of its descendants is in Z, or

   2. ω is not a collider on U and ω is in Z.

   If all such paths are blocked then X and Y are d-separated by Z. If the variable sets X and Y are d-separated by Z, they are independent conditional on Z in all probability distributions such a graph can represent.

7. Remark

   (1). In fig(3.8)(a), in path a-b-d-e, d is not a collide here.

   (2). In fig(3.8)(a), in path a-b-d-c, d is A collide here. Why a-b-d-c is A path from a to c? Because there are nodes and edges link to each other, information could pass via this path. It's unnecessary to consider the direct of arrows.