Conditional Independence in Bayesian Network

Kai Zhang
kz298@cornell.edu

March 20, 2018

Problem

A Bayesian network is a DAG $G = (V, E)$. Suppose a subset $S \subset V, |\bar{S}| \geq 2$ is observed, our problem is that for any given $v_1, v_2 \in \bar{S}, v_1 \neq v_2$, check if $v_1 \perp v_2 \mid S$ holds.

Solution

First we define a subgraph graph $G_{\bar{S}} \subset G$ consisting of vertices in $\bar{S}$. Then we define 5 sets of vertices as below.

$$\begin{aligned}
\Phi_1 &= \left\{v \in \bar{S} : v_1 \rightsquigarrow v \text{ in } G_{\bar{S}}\right\} \\
\Phi_2 &= \left\{v \in \bar{S} : v_2 \rightsquigarrow v \text{ in } G_{\bar{S}}\right\} \\
\Phi_3 &= \left\{v \in \bar{S} : v \rightsquigarrow v_1 \text{ in } G_{\bar{S}}\right\} \\
\Phi_4 &= \left\{v \in \bar{S} : v \rightsquigarrow v_2 \text{ in } G_{\bar{S}}\right\} \\
\Phi_5 &= \left\{v \in \bar{S} : \exists v' \in S, v \rightsquigarrow v' \text{ in } G\right\} \cup S
\end{aligned}$$

, where '$v_1 \rightsquigarrow v$ in $G_{\bar{S}}$' means there exists a path from v_1 to v in $G_{\bar{S}}$; same rule applies to the others.

Then we conclude that $v_1 \perp v_2 \mid S$ i.f.f. all the following three conditions hold.

1) $v_2 \notin \Phi_1$ and $v_1 \notin \Phi_2$

2) $\Phi_3 \cap \Phi_4 = \phi$

3) either v_1 or v_2 is disconnected with Φ_5 in G

The correctness of the above solution is actually quite straightforward. The only thing we should focus on is that we should regard S as a multi-dimensional random variable so that we don't need to worry about the structure inside S.