Bayes Ball algorithm

Daniel Hsu (COMS 4721 Spring 2018)

The problem

Let $G = (V, E)$ be a directed acyclic graph, where the vertices $V = \{1, \ldots, n\}$ index a collection of random variables $X_1, \ldots, X_n$. The directed graphical model associated with $G$ is the family of distributions that factorize over $G$. Consider any $s, t \in V$ and $E \subseteq V \setminus \{s, t\}$. The problem is to determine if $X_s \perp \perp X_t | X_E$ holds for all distributions in the model (where $X_E := (X_v : v \in E)$), or if there exists some distribution in the model such that $X_s \not\perp \not\perp X_t | X_E$.

The algorithm

The Bayes Ball algorithm is a reachability algorithm for solving this problem. Starting from vertex $s$, we traverse the undirected “backbone” of $G$ (i.e., the undirected version of $G$ in which we ignore the orientation of edges) in attempt to reach vertex $t$. The rules for assessing the validity of paths are given below. If there is a valid path from $s$ to $t$, then we declare $X_s \not\perp \not\perp X_t | X_E$ for some distribution in the model. If there are no valid paths from $s$ to $t$, then we declare $X_s \perp \perp X_t | X_E$ for all distributions in the model.

Rules of Bayes Ball

The following rules are used to assess the validity of a candidate path $v_1, \ldots, v_\ell$ in the undirected “backbone” of $G$ (so either $(v_i, v_{i+1}) \in E$ or $(v_{i+1}, v_i) \in E$ for $i = 1, \ldots, \ell - 1$).

- **Rule 1**: Suppose a candidate path contains a sequence of three vertices $u, v, w$ that appear as
  \[ u \leftarrow v \rightarrow w \]
  when the edge orientations from $G$ are considered. If $v \in E$, then the path is invalid.

- **Rule 2**: Suppose a candidate path contains a sequence of three vertices $u, v, w$ that appear as
  \[ u \rightarrow v \rightarrow w \quad \text{or} \quad u \leftarrow v \leftarrow w \]
  when the edge orientations from $G$ are considered. If $v \in E$, then the path is invalid.

- **Rule 3**: Suppose a candidate path contains a sequence of three vertices $u, v, w$ that appear as
  \[ u \rightarrow v \leftarrow w \]
  when the edge orientations from $G$ are considered. If $v \notin E$, then the path is invalid.

Note that it is possible to have $u = w$ in Rules 1 and 3. (Rule 2 is never applicable when $u = w$ because $G$ is acyclic.)

If a candidate path is not declared invalid by any of these three rules, then it is valid.

Examples

We now use the Bayes Ball algorithm to check various conditional independence relations in directed graphical models.
Example 1

Consider the directed graphical model shown below.

To check: $X_1 \perp X_6 \mid X_{\{2,3\}}$.

Yes, this conditional independence relation holds for every distribution in the model.

To see this, observe that every path from 1 to 6 has to go through either 2 or 3. Consider the path 1, 2, 6. This appears as $1 \to 2 \to 6$,

but we are conditioning on $X_2$ (i.e., $2 \in E$). So by Rule 1, this path is invalid. Now consider the path 1, 3, 5, 6.
The first three vertices in this path 1, 3, 5 appears as

$1 \to 3 \to 5$.

But, again, we are conditioning on $X_3$, so by Rule 1, this path is invalid.

Example 2

Consider the same directed graphical model as in Example 1.

To check: $X_2 \perp X_3 \mid X_{\{1,6\}}$.

No, this conditional independence relation does not hold for some distribution in the model.

Consider the path 2, 6, 5, 3. The first sequence of three vertices in this path appears as

$2 \to 6 \leftarrow 5$.

Since we are conditioning on $X_6$ (so $6 \in E$), Rule 3 does not invalidate the path. The next sequence of three vertices in this path appears as

$6 \leftarrow 5 \leftarrow 3$.

Since we are not conditioning on $X_5$, Rule 2 does not invalidate the path. So the entire path 2, 6, 5, 3 is valid.

Example 3

Consider the directed graphical model shown below.
To check: $X_1 \perp X_2 \mid X_4$.

No, this conditional independence relation does not hold for some distribution in the model.

Consider the path $1, 3, 4, 3, 2$, which appears as

$$1 \rightarrow 3 \rightarrow 4 \leftarrow 3 \leftarrow 2.$$ 

It can be checked that this path is valid.

It is worth pointing out that the path $1, 3, 2$ is not valid. Indeed, this path appears as

$$1 \rightarrow 3 \leftarrow 2,$$

but we are not conditioning on $X_3$. So Rule 3 invalidates this path.

Nevertheless, there is at least one valid path from $1$ to $2$, which is all that is needed to establish that the conditional independence relation does not hold for some distribution in the model.