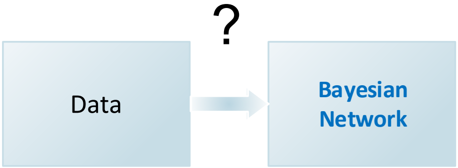
Learning the Structure of Bayesian Networks

Xiao Lin

Department of Computer Science and Engineering

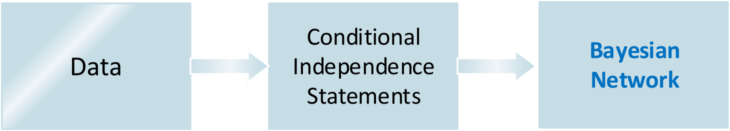
April 24, 2014

- Assume that you are given a bunch of cases generated by some unknown Bayesian network N over the universe $\mathcal{U}$ and you want to reconstruct the Bayesian network. What you will do is to learn the structure of the Bayesian network from the cases.



Constraint Based Learning Methods

- The constraint based methods establish a set conditional independence statements holding for the data, and use this set to build a network with d-separation properties corresponding to the conditional independence properties determined [Jensen and Nielsen, 2007].

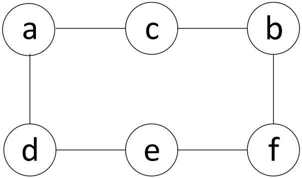


Notation for Conditional Independence

- I(a,b,χ): a is independent from b given χ
- I(a,b): shorthand for I(a,b,ϕ)
- I(a,b,c): a is independent from b given c

Notation for PC Algorithm

- $A_C ab$: the set of nodes adjacent to a or to b in gragh C, except for a and b themselves.
- $U_C ab$: the set of nodes in graph C on (acyclic) undirected paths between a and b, except for a and b themselves.



$$A_C ab = \{c, d, f\}$$
$$U_C ab = \{c, d, e, f\}$$
$$A_C ab \cap U_C ab = \{c, d, f\}$$

PC Algorithm [Spirtes and Glymour, 1991]

1. From the complete undirected graph C on the nodes set V.
2. $i = 0$.

   repeat

   - For each pair of nodes (a, b) adjacent in C, if $A_C ab \cap U_C ab$ has cardinality greater than or equal to i and a, b are independent conditional on any subsets of $A_C ab \cap U_C ab$ of cardinality less than i, delete a-b from C.
   - $i = i + 1$

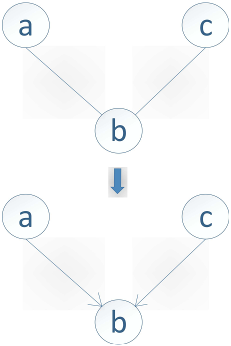
   until for each pair of adjacent nodes a, b, $A_C ab \cap U_C ab$ is of cardinality less than i.

   Call the resulting undirected graph F.
3. For each triple of nodes (a, b, c) such that the pair (a, b) and the pair (b, c) are each adjacent in F but the pair (a,c) are not adjacent in F, orient a-b-c as $a \to b \leftarrow c$ if and only if a and c are dependent on every subset of $A_F ab \cap U_F ab$ containing b. Output all graphs consistent with these orientations.

learning skeleton of BN

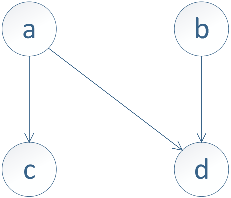
- $i = 0$, $I(a, b)$? If yes, remove the link (a, b)
- $i = 1$, $I(a, b, \{x\})$? If yes, remove the link (a, b).
  $\{x\}$ is any subset of $A_D ab \cap U_D ab$ with one node.
- $i = 2$, $I(a, b, \{x, y\})$? If yes, remove the link (a, b).
  $\{x, y\}$ is any subset of $A_F ab \cap U_F ab$ with two nodes.
- ... (until the cardinality of $A_F ab \cap U_F ab$ is less than i.)

orienting links



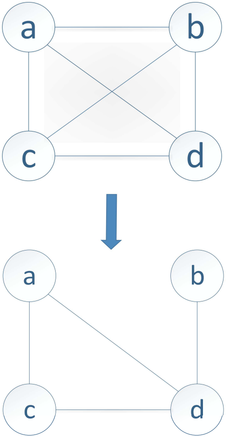
If and only if for every subset S of $A_C ac \cap U_C ac$ containing b, I(a, c, S) $\rightarrow$ Yes.

example



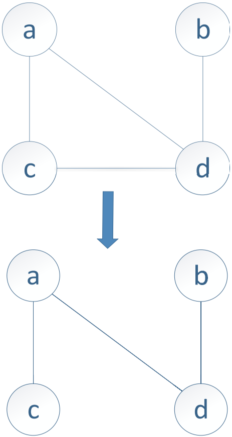
I(a, b)? Yes
I(a, c)? No
I(a, d)? No
I(b, c)? Yes
I(b, d)? No
I(c, d)? No
I(a, c, d)? No
I(a, d, c)? No
I(c, d, a)? Yes

example



$i = 0$
I(a, b)? Yes. remove (a, b)
I(a, c)? No
I(a, d)? No
I(b, c)? Yes. remove (b, c)
I(b, d)? No
I(c, d)? No

example



i = 1

$A_C ac \cap U_C ac = \{d\}$

I(a, c, d)? No

$A_C ad \cap U_C ad = \{c\}$

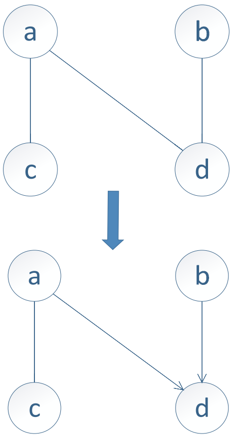
I(a, d, c)? No

$A_C cd \cap U_C cd = \{a\}$

I(c, d, a)? Yes. remove (c, d)

$A_C bd \cap U_C bd = \phi$

example



$A_C ab \cap U_C ab = \{\text{d}\}$
I(a, b, d)? Yes $\rightarrow$ converging connection

- Avoid new converging connection
- Avoid directed cycles

References

Finn V.Jensen and Thomas D.Nielsen (2007)
Bayesian Networks and Decision Graphs Springer: NY, USA, 2007; 230 -236.

Peter Spirtes and Clark Glymour (1991)
An Algorithm for Fast Recovery of Sparse Causal Graphs
Social Science Computer Review 1991, Vol. 9, No.1, 62-72.