Journal of Computational Science &Engineering 38 (2019) 999-1002

Available at http://www.asocse.org ISSN 1710-4068

Characterization of Maximal Ancestral Graph under Interventions

Qianwen Zang*, Cheng Xu, Rui Wang

Department of Mathematics and Statistics, Qingdao University, Qingdao 266071, China

Abstract: **Maximal ancestral graph provide a class of graph that can encode conditional independence relation that arise in DAG models with latent and selection variables. In this paper, we give an algorithm to constructing maximal ancestral graph given directed acyclic graph. In particular, we are interested in predicting the probability distribution of some random variables that would result under interventions, we show the change of graph structure if some variables were forced to take certain value.**

Keywords: Maximal ancestral graph; intervention; directed acyclic graph

Received 3 September 2018, Revised 7 October 2018, Accepted 10 October 2018

*Corresponding Author: Qianwen Zang, zqw491@163.com

1. Introduction

Graphical Markov models is used to represent the relationships between variables in probability in recent years, with nodes of graph corresponding to random variables and edges representing dependencies. Several classes of graphs with various independence interpretations have been described in literature. In particular, directed acyclic graphs(DAG) are widely used in many fields to model causal relationships and independence between random variables in multivariate distribution[1, 2].

Directed acyclic graph over a set of attributes or random variables which are fully given can represent causal structure. Many authors have characterized the common aspects of these explanations into one representation. Geiger et al[3] research the independence of Bayesian networks, Meek C[4] proves the completeness and faithfulness of Bayesian networks to probability distribution. In many cases, however, the data generating process might involve unobserved confounders or selection variables, and we need to consider DAGs with latent variables to model the process. There are many latent variables in real life. For example, in the economic field, financial analysts combine theoretical knowledge with "market awareness experience" to explain the current stock market changes and predict the future trend of the stock market; in the medical field, the recovery of patients after treatment is closely related to their own health lever (see Figure 1).

There are the following two types of problems: (1) hybrid problem: the common cause L of variables X and Y is unobserved, or worse, unobservable, which is called latent (hidden) variables; (2) selection bias: the common result S of variables X and Y affects sampling mechanism, for example, only one sample satisfying S=s is extracted, and the distribution of S is not observed, which is called selection variables. These problems may lead to false connections between observed variables.

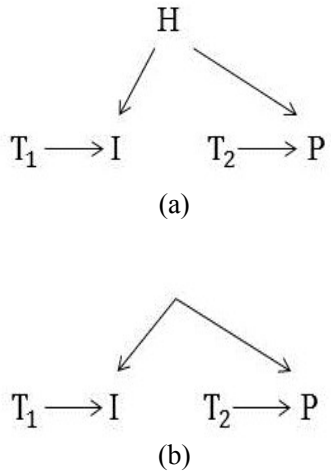

Figure 1. (a) A DAG with latent variable H; (b) The ancestral graph resulting from marginalizing H.

Consider the DAG shown in Figure 1(a) in which T_1, T_2, I, and P are observed variables, while H is unobserved. T_1 and T_2 represent treatments given to influenza patients. I is an opportunistic infection, and P is a measure of disease progression. As shown in Figure 1, the graph includes a bi-directed edge after marginalization. In this sense, the DAGs are not stable under marginalization, DAGs can no longer encode conditional independence and causal relationships. To overcome the limitation of DAG model, Richardson[5] first proposes the concept of ancestral graph, however, ancestral graph do not satisfy pairwise Markov property. Thus, we study maximal ancestral graph rather than ancestral graph. The statistical models associated with maximal ancestral graph retain many of the desirable properties that are associated with DAG model. Ali[6], Zhang J[7], Perković E et al[8], Evans R J[9] give some Markov properties of ancestral graph.

In particular, we are often interested in predicting

2018@ The American Computational Science Society. All rights reserved.

the probability distribution of some random variables that would result if some other variables were forced to take certain values, that is interventions. Characterization of DAGs under interventions was introduced by Hauser A et al[10].

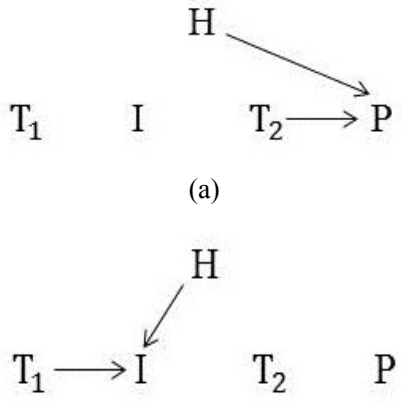

Figure 2. (a) intervention graph $D^{(\{I\})}$ under an intervention on I in Figure 1(a); (b) intervention graph $D^{(\{P\})}$ under an intervention on P in Figure 1(a).

In this paper, we have two main results. First, we propose a transition method from a DAG to a MAG in Section 2, and introduce the changes of graph structure of maximal ancestral graph under interventions in Section 3. In Section 4, a few concluding remarks are mentioned.

2. Preliminaries

G=(V, E) is a graph corresponding to P, V is a set of non-empty vertices, representing a group of random variables, E is a set of edges. For $a,b \in V$, $(a,b) \in E$ represent that there is an edge between a and b, we call a and b is adjacent. A mixed graph contain undirected(—), directed(→) and bi-directed edges(↔), the mixed graph studied in this paper is acyclic.

We use follow terminology to describe relations between vertices in a graph G:

$$\text{If} \begin{Bmatrix} a \rightarrow b \\ a \leftarrow b \\ a \leftrightarrow b \end{Bmatrix} \text{ in G then a is a } \begin{Bmatrix} \text{parent} \\ \text{child} \\ \text{spouse} \end{Bmatrix} \text{ of b and } \begin{Bmatrix} a \in pa(b) \\ a \in ch(b) \\ a \in sp(b) \end{Bmatrix}$$

An independence model over a set V is a set I of ternary relations $\langle X,Y|Z\rangle$ where X, Y and Z are disjoint subsets of V, and X and Y is not empty. The interpretation of $\langle X,Y|Z\rangle \in I$ is that X and Y are independent given Z.

A path p=<a,v_1,v_2,…,v_k,b> between a and b in G is a sequence of distinct vertices, a and b are endpoints, v_i is non-endpoint (interior point), if every edge between a and b is of the form →, with the arrowheads pointing toward b, p is a directed path from a to b, then a is an ancestor of b, we use $a \in an(b)$ to denote it.

There is an edge connection between a and b, we state that a and b is adjacent node, is denoted by a~b, otherwise a and b are non-adjacent.

$a \rightarrow b \leftarrow c$ is a V-structure, and denoted by V_s, b is a collider, otherwise is non-collider.

Two paths (including V_s) are called endpoint-identical if presence or lack of arrowheads pointing to endpoints on the paths are the same in both, for example, $a \rightarrow b \rightarrow c$ and $a \rightarrow b \leftrightarrow c$ are endpoint-identical.

A mixed graph is ancestral if:

(i) there are no directed cycle;

(ii) if an edge $a \leftrightarrow b$ is in graph then a is not an ancestor of b and b is not an ancestor of a;

(iii) if there is an undirected edge between a and b, then a and b have no spouses or parents.

By definition of ancestral graph, we can conclude that directed acyclic graphs and undirected graphs are special cases of ancestral graphs. Directed acyclic graph models have separation criteria which is d-separation, and corresponding ancestral graphs also have separation criteria to define conditional independence relations.

In an ancestral graph, a path p between a and b is said to be m-connecting given Z if the following hold:

(i) no non-collider on p is in Z;

(ii) every collider on p is an ancestor of vertex in Z.

Two vertices a and b are said to be m-separated given Z if there is no path m-connecting a and b in G.

An inducing path p between vertices a and b in an ancestral graph is a path on which every non-endpoint is both a collider on p and an ancestor of at least one of the endpoint a and b.

An ancestral graph is maximal, if for every pair of nonadjacent vertices (a,b), there is a set Z such that a and b are m-separated by Z.

Next we give a set of orientation rules to construct MAG M given DAG D.

Step 1. For latent variable h which is neither endpoint nor collider, add endpoint-identical edge at h, then delete the original edge, generate bi-directed edges in D;

Step 2. Construct a undirected graph M with the same adjacencies as G; and add directed edges to M;

Step 3. For bi-directed edges in D, orient the edge as $a \rightarrow b$ in M if a is an ancestor of b in D, orient the edge as $a \leftarrow b$ in M if b is an ancestor of a in D, otherwise copy $a \leftrightarrow b$ to M;

Step 4. For non-adjacent vertices a and b which exist an inducing path between them, orient the edge as a→b in M if a is an ancestor of b in D, orient the edge as a←b in M if b is an ancestor of a in D, otherwise copy a↔b to M.

Maximal ancestral graph do satisfy pairwise Markov

2018@ The American Computational Science Society. All rights reserved.

property: two non-adjacent vertices a and b, there is a set Z, such that a and b are m-separated by Z.(see Figure 3(d))

The process of non-adjacent nodes with an inducing path among Step 4 is a directed consequence of fact that AGs do not satisfy a pairwise Markov property. Obtain a MAG by adding edges to this type of non-adjacent nodes.

An example of constructing corresponding maximal ancestral graph given acyclic directed graph is as follows, V_8, V_9, V_{10} is latent variables.

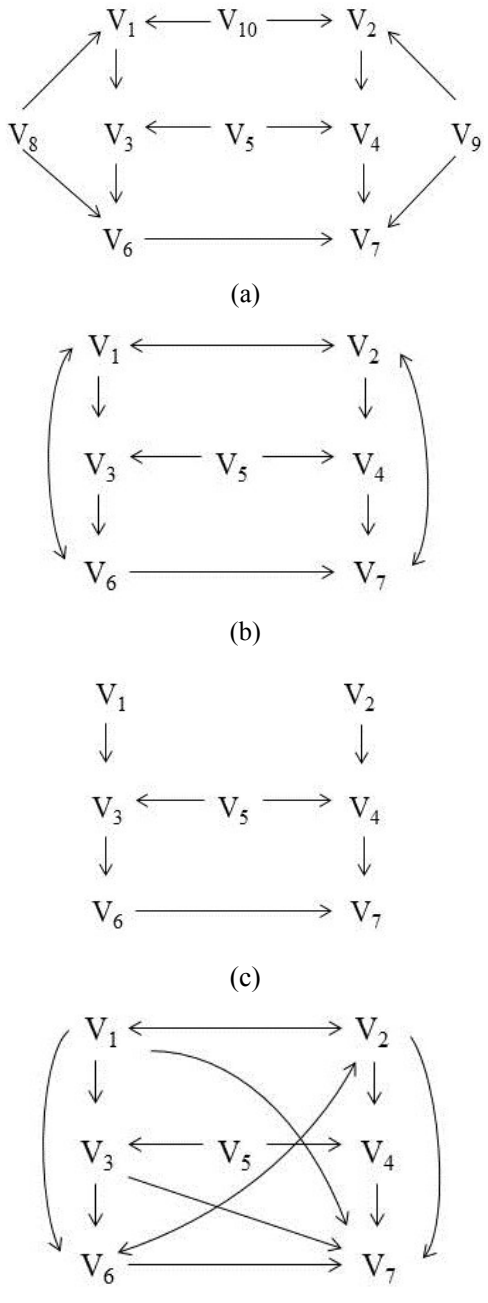

Figure 3. (a) A DAG with latent variables; (b) obtained graph after Step 1; (c) the graph after Step 2; (d) The MAG after Step 3-4.

3. MAG under interventions

As illustrated in Figure 2, when we intervene X =x in directed acyclic graph, only need to delete the edges between X and its parents. We are also in interested in intervention of maximal ancestral graph.

An edge is called into (or out of) a vertex if mark of the edge at the vertex is an arrowhead (or tail), for example, there is two types of into a: $a \leftarrow b$, $a \leftrightarrow b$.

A directed edge $a \rightarrow b$ in MAG is visible if there is a vertex c which is not adjacent to b, such that either there is an edge between c and a that is into a, or there is a collider path between a and c that is into a and every node on the path is a parent of b. Otherwise $a \rightarrow b$ is invisible.

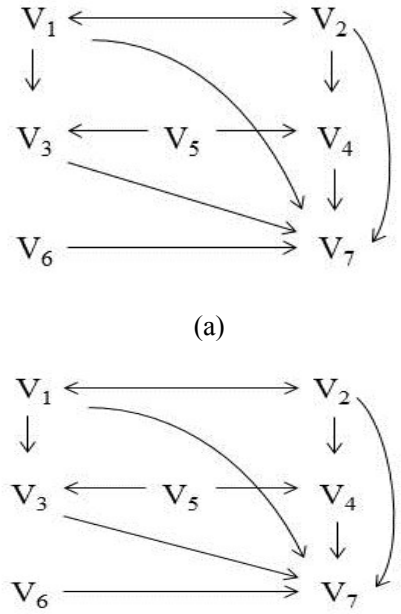

Figure 4. (a) obtained graph intervention V_6 in Figure 3(d); (b) obtained graph intervention V_2 in Figure 3(d).

We have known a fact that if $a \rightarrow b$ is visible in MAG, for every DAG D corresponding the MAG M, there is no inducing path between a and b relative to the set of latent variables that is also out of b. Then for this DAG, the intervention graph resulting from deleting edges into b in D, thus, there is not any inducing path between a and b relative to latent variables. According to the pairwise Markov property of MAGs: there is no inducing path between any two nonadjacent vertices, MAG will not contain any edge between a and b. So deleting visible arrows into b in MAGs corresponds to eliminating edge into b in DAGs.

For invisible arrow $a \rightarrow b$ and bi-directed edge

2018@ The American Computational Science Society. All rights reserved.

$a \leftrightarrow b$ in MAG, this implies that there is a hidden variable between a and b such that $a \rightarrow h \leftarrow b$ in underlying DAG model. So deleting edges into b in DAG corresponds to invisible arrow $a \rightarrow b$ and $a \leftrightarrow b$ disappear in MAG.

In a word, when we intervene X=x, eliminating edges into X in underlying DAGs represented by MAG corresponds to deleting edges into X in MAGs. (Figure 4)

4. Conclusion

In this paper, we present a set of orientation rules that construct corresponding maximal ancestral graph given directed acyclic graph. And we propose graph structure of maximal ancestral graph under interventions: delete all those edge in MAGs that are into variables, which are forced to take certain values.

Different maximal graph can encode same conditional independence, which is called Markov equivalence. So discuss interventional Markov equivalence is probably a worthy project.

References

[1] Lauritzen S L, Clarendon: Oxford University Press, 1996.
[2] Neapolitan R E, Upper Saddle River:Prentice Hall, 2004.
[3] Geiger D, Networks, 2010, 20(5):507-534.
[4] Meek C. Proc. Conf. on Uncertainty in Artificial Intelligence (UAI-95), 2013:411-418.
[5] Richardson T, Spirtes P, Annals of Statistics, 2002, 30(4):962-1030.
[6] Ali R A, Annals of Statistics, 2009, 37(5B):2808-2837.
[7] Zhang J, Journal of Machine Learning Research, 2008, 9(3):1437-1474.
[8] Perković E, Textor J, Kalisch M, et al, Journal of Machine Learning Research, 2018, 18(220):1-62.
[9] Evans R J, Scandinavian Journal of Statistics, 2016, 43(3):625-648.
[10] Hauser A, Bühlmann, Peter, Journal of Machine Learning Research, 2012, 13(1):2409-2464.

2018@ The American Computational Science Society. All rights reserved.