TECHNICAL REPORT
R-392
April 2012

A note on the pairwise Markov condition in directed Markov fields

Judea Pearl
University of California, Los Angeles
Computer Science Department
Los Angeles, CA, 90095-1596, USA
(310) 825-3243
judea@cs.ucla.edu

April 9, 2012

Abstract

It is well known that, in directed Markov fields, the pairwise Markov condition does not imply the global Markov condition, unless the distribution is strictly positive. We introduce a stronger version of the pairwise condition which requires that every nonadjacent pair be independent conditional on *every* set that separates the pair in the graph. We show that this stronger condition is equivalent to the global Markov condition (for all probability distributions.) We generalize this result to abstract dependency models, and show that a weaker condition holds for compositional graphoids.

1 Introduction

A probability distribution is said to be Markov relative to a directed acyclic graph (DAG) G if every d-separation condition in G is confirmed by a corresponding conditional independence condition in P. This condition is also called "globally Markov" since it applies to all subsets of variables and, when it holds, we say that P and G are "compatible" (Pearl, 1988; Lauritzen, 1996).

Several local conditions have been devised which apply to singleton variables, among them the "local Markov condition" and the "pairwise Markov condition." The local Markov condition requires that every variable be conditional independent of its non descendants, given its parents. This condition is often taken as a definition of Bayesian Networks and it can be shown to be equivalent to the global Markov condition for all probability distributions Pearl and Verma (1987); Geiger et al. (1990); Lauritzen (1996).[1] The pairwise Markov condition requires that every pair (x, y) of non adjacent variables with y non-descendant of

[1] This global-local equivalence holds in fact in all dependency models that obey the semi-graphoid axioms, not necessarily probabilistic.

x be independent, conditional on all other non-descendants of x (Lauritzen, 1996, p. 50). This condition, however, while entailed by the global Markov condition, is too weak to ensure global compatibility unless the distribution is strictly positive (a simple counterexample is given below).

This idiosyncratic feature of pairwise independencies sets directed Markov fields apart from their undirected counterparts, in which the global, local and pairwise conditions coincide Pearl and Paz 1986; Pearl 1988, p. 98; Lauritzen 1996. This note introduces a stronger version of the pairwise condition, named *pairwise compatibility*, which requires that every nonadjacent pair be independent conditional on *all* its separating sets in the DAG. We show that this stronger condition is equivalent to the global Markov condition for all probability distributions, including those that impose logical or equality constraints.

Notation

- Let P be a joint distribution function on a set V of variables, and G a Directed Acyclic Graph (DAG) with vertices corresponding to variables in V.
- Let X, Y, and Z stand for sets of variables in V, and x, y, z singleton variables in V.
- Let $(X, Y, Z)_G$ stand for the assertion "Y d-separates X from Z in G."
- Let $(X, Y, Z)_P$ stand for the assertion: "Y is independent of X given Z in P."

Definitions

- A DAG G is said to be set-compatible (SP) with probability function P iff
$$(X, Y, Z)_G \implies (X, Y, Z)_P \tag{1}$$
for every three sets X, Y, and Z in V.
- A DAG G is said to be *pairwise-compatible* (PC) with probability P iff
$$(x, Y, z)_G \implies (x, Y, z)_P \tag{2}$$
for every set Y and every pair of singleton variables x and z.

Clearly, set-compatibility implies pairwise compatibility. We will show that the converse is also true.

Theorem 1 *Pairwise-compatibility implies set-compatibility*

Proof:
It is enough to prove that pairwise compatibility implies $(x, pa(x), nd(x))_P$, where $pa(x)$ is the set of parents of x, and $nd(x)$ is the set of nondescendants of x. The reason is that the condition $(x, pa(x), nd(x))_P$ for all x implies set-compatibility (See *Causality*, Pearl 2009, p. 19, Theorem 1.2.7).
Proof by induction:
Assuming PC, let us prove that $(x, pa(x), nd(x))_P$ holds for every variable x, using induction on the cardinality of sets in $nd(x)$.

Let I_k stand for the hypothesis:

I_k: if $(x, Y, z)_G \implies (x, Y, z)_P$ for every set Y and every singletons x and z, then, for every x we have: $(x, pa(x), S_k)_P$, where S_k is any set of nondescendants of x, such that $card(S_k) \leq k$.

Base: For $k = 1$, I_k is trivially true, because I_1 reduces to an identity. Indeed, I_1 says: if $(x, Y, z)_G \implies (x, Y, z)_P$ for every set Y and every singletons x and z, then, for every x we have: $(x, pa(x), S_1)_P$, where S_1 is any set of nondescendants of x, such that $card(S_1) \leq 1$. But the sets S_1 are all singletons, hence whatever is true for z, would be true for S_1 as well.

The induction step is: $I_k \Rightarrow I_{k+1}$

Let S_{k+1} be any set of nondescendants of x, such that $card(S) \leq k + 1$.

Let z be any singleton variable in S_{k+1}, and $S = S_{k+1} \backslash z$

Let T stands for the claim of pairwise compatibility, i.e.,:

$$(x, Y, z)_G \implies (x, Y, z)_P \text{ for every singletons } x \text{ and } z,$$

From I_k we have $T \Rightarrow (x, pa(x), S)_P$, because the cardinality of S is not exceeding k. We also have $(x, pa(x), z, S)_G$ because S and z are both in $nd(x)$ of G. This implies $(x, pa(x), z, S)_G$, because d-separation is a graphoid. and any graphoid obeys the weak union axiom:

$$(x, pa(x), zS)_G \Rightarrow (x, pa(x)z, S)_G,$$

Thus, $T \Rightarrow (x, pa(x)z, S)_P$, because S is of card $\leq k$. and I_k is assumed to hold.

In summary, We now have

$$T \Rightarrow (x, pa(x), S)_P,$$

and

$$T \Rightarrow (x, pa(x)z, S)_P,$$

from which we conclude, using the contraction axiom for probabilistic independence (*Causality*, Pearl 2009, p. 11):

$$T \Rightarrow (x, pa(x), Sz)_P,$$

or

$$T \Rightarrow (x, pa(x), S_{k+1})_P,$$

This completes the inductive step QED

Example 1 *Let $x = y = z$ and let w be independent of x with*

$$P(x = 1) = P(x = 0) = P(w = 1) = P(w = 0) = 1/2$$

This distribution is pair-wise Markov with respect to the graph G_1 in Fig. 1(a), because every non adjunct pair of variables x, y with y nondescendant of x is independent conditional on all other nondecendants of x (in our case $\{z, w\}$). However, this distribution is not set-compatible with the graph, because $(x, w, yz)_G$ holds in G_1 while $(x, w, yz)_P$ does not hold in P. Now examine our strong pair-wise condition, PC; (x, y) are nonadjacent in G_1, with two separating sets, $\{w\}$ and $\{wz\}$. PC requires not only $(x, wz, y)_P$ but also $(x, w, y)_P$. While the first requirement holds in P, the second does not, which renders G_1 and P pair-wise incompatible.

In contrast, the graph G_2 depicted in Fig. 1(b) is pairwise compatible with P. Here, the only adjacent pair is (x, z) and it is separated by one set $\{wy\}$. Accordingly, pairwise comparability requires that x and z be independent conditional on $\{zw\}$, a requirement satisfied in P.

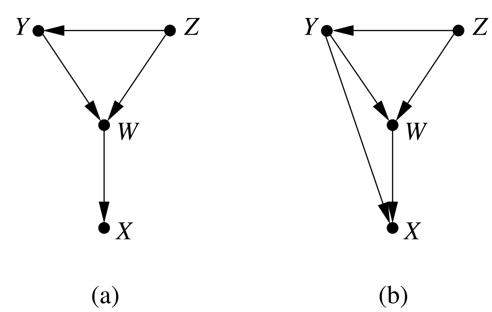

Figure 1:

2 Conclusions

We showed that a strong version of the pairwise condition is sufficient to guarantee the global Markov condition in directed Markov fields. The strong version requires that every pair of nonadjacent nodes be independent conditional on every set that d-separates the pair. Clearly the ability to replace set-compatibility with pairwise compatibility is not unique to graphs and probabilities, but extends to dependency models for which the semi-graphoid axioms hold.

Acknowledgments

This research was supported in parts by grants from NIH #1R01 LM009961-01, NSF #IIS-0914211 and #IIS-1018922, and ONR #N000-14-09-1-0665 and #N00014-10-1-0933.

References

Geiger, D., Paz, A. and Pearl, J. (1990). Learning simple causal structures. *Journal of Intelligent Systems* **50** 510–530.

Lauritzen, S. (1996). *Graphical Models*. Clarendon Press, Oxford.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, San Mateo, CA.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. 2nd ed. Cambridge University Press, New York.

Pearl, J. and Paz, A. (1986). On the logic of representing dependencies by graphs. In *Proceedings 1986 Canadian AI Conference*. Montreal, Ontario, Canada.

Pearl, J. and Verma, T. (1987). The logic of representing dependencies by directed acyclic graphs. In *Proceedings of the Sixth National Conference on AI (AAAI-87)*. Seattle, WA.