J. Appl. Prob. **29**, 745–749 (1992)

ON EQUIVALENCE OF MARKOV PROPERTIES OVER UNDIRECTED GRAPHS

F. MATÚŠ [1]

Institute of Information Theory and Automation, Prague

Abstract

The dependence of coincidence of the global, local and pairwise Markov properties on the underlying undirected graph is examined. The pairs of these properties are found to be equivalent for graphs with some small exluded subgraphs. Probabilistic representations of the corresponding conditional independence structures are discussed.

MARKOV RANDOM FIELDS, MARKOV PROPERTIES, CONDITIONAL INDEPENDENCE

1. Introduction. Conditional independence restrictions of the Markov type together with factorizations of probability distributions of Markov fields have been investigated over two decades. It is well known that if the distribution factorizes with respect to an undirected graph then it has the corresponding global Markov property which implies the local Markov property yet stronger than the pairwise one [see Lauritzen (1989),Whittaker (1990)].

Alternatively, under a positivity assumption on the density of the distribution a famous result, ascribed to Hammersley and Clifford, asserts the equivalence of the three Markov properties to the admissibility of factorizations [for an account of references see Speed (1979)]. In general, if the assumption is dropped out the assertion fails. Let us remark that for the directed graphs a similar statement does not require a positivity [Lauritzen, Dawid, Larsen and Leimer (1990)].

Considering a special fixed graph the global Markov property related to it could be still equivalent to the local one or even to the pairwise one. The purpose of the present note is to determine the classes of undirected graphs with respect to which pairs of these properties coincide. Our solution is based on a reformulation of their definitions emphasizing graphical conditional independence structures. At the end we discuss briefly probabilistic representations of these structures.

---

[1]Postal address: Institute of Information Theory and Automation, Czechoslovak Academy of Sciences, Pod vodárenskou věží 4, 182 08 Prague, Czechoslovakia.

This research was supported partially by Internal Grant 27510 of Czechoslovak Academy of Sciences.

AMS 1980 subject classifications. Primary 68H, secondary 05C.

2. Graphtheoretic preliminaries. Let N be a finite nonempty set and $\mathcal{P}_2(N)$ the class of all two-element subsets of N. Elements i, j, k, l of N will be sometimes considered for singletons of N. The union of subsets I, J of N is written as IJ. The class of all ordered pairs $(ij|K)$ [we prefer the nonstandard deliminator $|$], where $ij \in \mathcal{P}_2(N)$ and $K \subset N - ij$, will be denoted by $\mathcal{R}(N)$.

We shall deal with simple undirected graphs $G = (N, E)$, $E \subset \mathcal{P}_2(N)$. The boundary ∂I of a subset I of vertices is the set $\partial I = \{j \in N - I;\ \exists i \in I,\ ij \in E\}$ and the closure $\bar{I}$ of I is $\bar{I} = I \cup \partial I$. A subset K of N is said to separate $I \subset N$ from $J \subset N$ if all chains from I to J intersect K, e.g. if the conditions $s \geq 1$, $i_1i_2, i_2i_3, ..., i_si_{s+1} \in E$, $i_1 \in I$ and $i_{s+1} \in J$ imply $(i_1i_2 \cdots i_{s+1}) \cap K \neq 0$.

The subgraph G_K of a graph $G = (N, E)$ generated by a nonempty subset K of N is the graph $G_K = (K, E \cap \mathcal{P}_2(K))$. Let us denote by $\mathcal{G}_\square$ the class of all graphs (N, E) which do not have a subgraph $(ijkl, \{ij, kl\})$ having four vertices and two nonadjacent edges; a graph (N, E) belongs to $\mathcal{G}_\square$ if and only if the dual graph $(N, \mathcal{P}_2(N) - E)$ does not have a cycle of length four without a chord, i.e. a subgraph of the form $\square$. Similarly, let $\mathcal{G}_{<\triangle}$ be the class of all graphs with forbidden subgraphs $(ijk, \{jk\})$ and $(ijk, \emptyset)$ of three vertices and at most one edge. The duals of these graphs do not have subgraphs of the form $<$ and $\triangle$, or equivalently, they do not have two adjacent edges.

Let us introduce these three subsets of $\mathcal{R}(N)$ associated with a graph $G = (N, E)$

$$\begin{aligned}
\mathcal{K}_g(G) &= \{(ij|K) \in \mathcal{R}(N);\ K \text{ separates } i \text{ from } j\} \\
\mathcal{K}_l(G) &= \{(ij|K) \in \mathcal{R}(N);\ \partial i \subset K \text{ or } \partial j \subset K\} \\
\mathcal{K}_p(G) &= \{(ij|N - ij);\ ij \in \mathcal{P}_2(N) - E\}
\end{aligned}$$

[the separation and the boundaries are considered with respect to the graph G]. Evidently, $\mathcal{K}_g(G) \supset \mathcal{K}_l(G) \supset \mathcal{K}_p(G)$.

LEMMA 1. $\mathcal{K}_g(G) = \mathcal{K}_l(G)$ if and only if G belongs to $\mathcal{G}_\square$.

PROOF. If G has a subgraph $(ijkl, \{ik, jl\})$ of four vertices then $K = N - ijkl$ separates i from j and at the same time $k \in \partial i - K, l \in \partial j - K$, thus $(ij|K) \in \mathcal{K}_g(G) - \mathcal{K}_l(G)$.

On the other hand, if $G \in \mathcal{G}_\square$ and $ij \notin E$ then for any $k \in \partial i - \partial j$ and $l \in \partial j - \partial i$ we have $kl \in E$. Now, if K separates i from j, $(ij|K) \in \mathcal{K}_g(G)$, then $K \supset \partial i \cap \partial j$ and at least one of the sets $\partial i - K$, $\partial j - K$ must be empty [otherwise for some $k \in \partial i - K$ and $l \in \partial j - K$ the chain ik, kl, lj is out of K], thus $(ij|K) \in \mathcal{K}_l(G)$. Q.E.D.

LEMMA 2. $K_l(G) = K_p(G)$ if and only if G belongs to $\mathcal{G}_{<\triangle}$.

PROOF. If $G \notin \mathcal{G}_{<\triangle}$ then there are $i, j, k \in N$ distinct such that $ij, ik \notin E$. But then $N - ijk \supset \partial i$ and hence $(ij|N - ijk) \in \mathcal{K}_l(G) - \mathcal{K}_p(G)$.

If $G \in \mathcal{G}_{<\triangle}$ and $ij \notin E$ then $\partial i = \partial j = N - ij$ [otherwise for some $k \in (N - ij) - \partial i$ or for some $k \in (N - ij) - \partial j$ the subgraph G_{ijk} has at most one edge and hence $G \notin \mathcal{G}_{<\triangle}$] and thus $(ij|K) \in \mathcal{K}_l(G)$ implies $(ij|K) \in \mathcal{K}_p(G)$. Q.E.D.

3. Conditional independence. Let $\xi = (\xi_i)_{i \in N}$ be a system of random variables indexed by N and let for $K \subset N$ $\xi_K = (\xi_k)_{k \in K}$ be the corresponding subsystem; $\xi_\emptyset$ is supposed to be a constant. The statement "ξ_I is conditionally independent of ξ_J given ξ_K", $I, J, K \subset N$, we shorten to $\xi : I \perp J | K$.

The following two equivalences are valid for any ξ and $I, J, K, L \subset N$

$$\xi : I \perp J|K \;\Leftrightarrow\; \xi : J \perp I|K,$$

$$\xi : I \perp JK|L \;\Leftrightarrow\; \xi : I \perp J|KL \text{ and } \xi : I \perp K|L$$

[see Dawid (1979), Mouchart and Rolin (1984), Whittaker(1990)]. Using them we get [cf. with Proposition 1, in Matúš (1991)]

LEMMA 3. For any ξ and $I, J, K \subset N$

$$[\xi : I \perp J|K] \;\Leftrightarrow\; [(\forall\, i\, \in I)\,(\forall\, j\, \in J)\,(\forall\, L \subset N)\;(K \subset L \subset IJK - ij \Rightarrow \xi : i \perp j|L)].$$

PROOF. The implication $\Rightarrow$ is a simple consequence of the two properties above. To prove the reverse we can assume that I and J are nonempty. If they are singletons i, j not contained in K the right hand side gives $\xi : i \perp j|K$; if i or j is in K the left hand side holds trivially. We proceed by induction on the sum of cardinalities of I and J. Using the symmetry of I and J we can decompose $J = J_1 J_2$ into two disjoint nonempty sets J_1 and J_2. By the induction argument we deduce from the right hand side $\xi : I \perp J_1|K$ and $\xi : I \perp J_2|J_1 K$. An application of the second property closes the induction step. Q.E.D.

We remark that there is a profusion of properties of the above type as the recent works Studený (1989) and Studený (1991) show. They are all not known even for four random variables.

4. The Markov properties. In this section we reformulate the definitions of Markov properties and analyse how they depend on the underlying graph. To this end the following notation will be employed. For a system ξ let $[\xi]$ be the set of those elements $(ij|K)$ of $\mathcal{R}(N)$ for which $\xi : i \perp j|K$.

DEFINITION. Let $G = (N, E)$ be a graph and $\xi = (\xi_i)_{i \in N}$ a system of random variables. The system ξ is said to obey the *global, local, pairwise Markov property* w.r.t. G if $\mathcal{K}_g(G)$, $\mathcal{K}_l(G)$, $\mathcal{K}_p(G)$ is contained in $[\xi]$, respectively.

The usual definition of these Markov properties [see Whittaker (1990)] coincides obviously with the above one for the pairwise Markov property. In the remaining two cases the more strict independences $\xi : I \perp J|K$ for K separating I from J [$I, J, K \subset N$ disjoint] and $\xi : i \perp N - \bar{i}|\partial i$ [$i \in N$] are demanded, respectively. But, if $K_g(G) \subset [\xi]$ then having a disjoint triple $I, J, K \subset N$, K separating I from J, we see that for any $i \in I$, $j \in J$ and L such that $K \subset L \subset N - ij$ the set L also separates i from j, thus $\xi : i \perp j|L$ and by Lemma 3 $\xi : I \perp J|K$. So, our definition of the global Markov property is equivalent to the standart one and a similar argumentation shows that this is valid for the local Markov property, too.

In a paper of Speed(1979) ξ is said to be *regionally Markov* w.r.t. G if ξ satisfies $\xi : I \perp N - I|\partial I$ or, what is the same, $\xi : I \perp N - \bar{I}|\partial I$, for any $I \subset N$. Since ∂I separates I from $N - \bar{I}$ this Markov property results from the global one. On contrary, if ξ is regionally Markov and if K separates I from J [$I, J, K \subset N$ disjoint], then K separates also some L from some M [$L, M \subset N$ disjoint], where $I \subset L$, $J \subset M$ and $LMK = N$. Hence, there is no edge between L and M and $\partial L \subset K$ follows. From $\xi : L \perp N - \bar{L}|\partial L$ we conclude $\xi : I \perp J|K$, i.e. ξ is globally Markov [the formulation in Speed (1979), p.193, above the section 4.3. is not precise].

PROPOSITION 1. The local and global Markov properties [of any system ξ of random variables] w.r.t. a graph G coincide if and only if G belongs to $\mathcal{G}_\Box$.

PROOF. If $G \in \mathcal{G}_\Box$ then $K_l(G) = K_g(G)$ by Lemma 1 and the Markov properties are equivalent by Definition.

If $G \notin \mathcal{G}_\Box$ then there are four distinct vertices $i, j, k, l \in N$ such that $ij, kl \notin E$, $k \in \partial i - \partial j$, $l \in \partial j - \partial i$ and $(ij|N - ijkl) \in K_g(G)$. Let η be a nonconstant random variable and let $\xi_m = \eta$ for $m \in ijkl$, let ξ_m be a constant otherwise and let $\xi = (\xi_m)_{m\in N}$. Then $\xi : i \perp j|N - ijkl$ is not valid and thus ξ is not globally Markov. At the same time $\xi : m \perp N - \bar{m}|\partial m$ holds trivially for $m \in N - ijkl$. For $m \in ijkl$ the local Markov property demands η to be independent of η given η and thus it is satisfied. Q.E.D.

PROPOSITION 2. The pairwise and local Markov properties w.r.t. a graph G are equivalent [for any ξ] if and only if $G \in \mathcal{G}_{<\triangle}$.

PROOF. One direction of the claim is settled by Lemma 2 similarly as in the previous proof. Conversely, if $G \notin G_{<\triangle}$ then there are ij, $ik \in \mathcal{P}_2(N) - E$, $j \neq k$. We set ξ_m to be a constant for $m \in N - ijk$ and $\xi_m = \eta$ otherwise as above. Again, the pair $(ij|N - ijk)$ belongs to $\mathcal{K}_l(G)$ and $\xi : i \perp j|N - ijk$ does not hold. Hence ξ is not locally Markov and it is easily seen to be pairwise Markov. Q.E.D.

COROLLARY. If the pairwise and local Markov properties w.r.t. a graph G are equivalent then they are equivalent to the global one as well [note that $\mathcal{G}_{<\triangle} \subset \mathcal{G}_\Box$].

5. Remarks. Let $\mathcal{K}$ be a subset of $\mathcal{R}(N)$ and let there exist a system ξ of random variables such that $\mathcal{K} = [\xi]$. To comment this situation we coin, starting from Matúš (1991), the phrase $\mathcal{K}$ *is probabilistically (p-) representable*; ξ is called *a p-representation of* $\mathcal{K}$. Propositions 1 and 2 will follow easily from Lemmas 1 and 2 and the fact that all conceivable classes $\mathcal{K}_l(G)$ and $\mathcal{K}_p(G)$ are p-representable.

It is easy to verify that $\mathcal{K}_g(G)$, $\mathcal{K}_l(G)$ and $\mathcal{K}_p(G)$ are ascending

$$(ij|K) \in \mathcal{K} \text{ and } K \subset L \subset N - ij \;\Rightarrow\; (ij|L) \in \mathcal{K}$$

and satisfy

$$(ij|kL) \in \mathcal{K} \text{ and } (ik|L) \in \mathcal{K} \quad \Rightarrow \quad (ij|L) \in \mathcal{K}$$

$$(ij|kL) \in \mathcal{K} \text{ and } (ij|lL) \in \mathcal{K} \text{ and } (kl|L) \in \mathcal{K} \quad \Rightarrow \quad (ij|L) \in \mathcal{K}.$$

In Matúš (1991) we proved that any $[\xi]$ [ξ taking finite number of values] which is ascending satisfies the two above implications and vice versa, and there the proof is rather intricate, that any $\mathcal{K} \subset \mathcal{R}(N)$ holding these three implications is p-representable. Hence, all $\mathcal{K}_g(G)$, $\mathcal{K}_l(G)$ and $K_p(G)$ are p-representable [for the case of the global Markov property see also Pearl (1988)].

At the end we indicate how one can construct p-representations of the local Markov conditional independence structures $\mathcal{K}_l(G)$. Let us denote by $\mathcal{M}$ the following family $E \cup \{ijkl;\ ij \in \mathcal{P}_2(N) - E,\ k \in \partial i,\ l \in \partial j\}$ of two-, three- and four-element subsets of N and let $(\eta^M)_{M \in \mathcal{M}}$ be a system of nonconstant mutually independent random variables. Every system $\xi^M = (\xi^M_m)_{m \in N}$, $M \in \mathcal{M}$, of random variables defined by $\xi^M_m = \eta^M$ if $m \in M$ and ξ^M_m equals a constant if $m \in N - M$ obey the local Markov property $[\xi^M] \supset \mathcal{K}_l(G)$. Then also $\xi = ((\xi^M_m)_{M \in \mathcal{M}})_{m \in N}$ is locally Markovian and we claim that this composed system of random variables is a p-representation of $\mathcal{K}_l(G)$. Indeed, if to every $(ij|K) \in \mathcal{R}(N) - \mathcal{K}_l(G)$ we take $M = ij$, if $ij \in E$, and $M = ijkl$, if $ij \in \mathcal{P}_2(N) - E$, $k \in \partial i - K \neq \emptyset$, $l \in \partial j - K \neq \emptyset$ [so that $M \in \mathcal{M}$], we obtain that $\xi^M : i \perp j|K$ is not valid. Consequently, $[\xi] = \mathcal{K}_l(G)$. This situation could be commented, in terms of Frydenberg (1990), by saying "the class $\{\xi^M;\ M \in \mathcal{M}\}$ is locally Markov perfect".

P-representations of the global and pairwise Markov conditional independence structures $\mathcal{K}_g(G)$, $\mathcal{K}_p(G)$ can be found quite analogously; it is necessary only to replace in the above construction the family $\mathcal{M}$ by the family of all connected sets of vertices of G and by $\mathcal{M} = E \cup \{ijk;\ i,j,k \in N \text{ distinct}\}$, respectively.

References

DAVID, A.P. (1979) Conditional independence in statistical theory (with discussion). *J. Roy. Statist. Soc. Ser. B* **41** 1-31.

FRYDENBERG, M.(1990) Marginalization and collapsibility in graphical interaction models. *Ann. Statist.* **18** 790-805.

LAURITZEN, S.L. (1989) *Lectures on Contingency Tables.* (3rd. Ed.) University of Aalborg Press: Aalborg.

LAURITZEN, S.L., DAWID, A.P., LARSEN, B.N., LEIMER, H.G. (1990) Independence properties of directed Markov fields. *Networks* **20** 579-605.

MATÚŠ, F. (1991) Ascending and descending conditional independence relations. In *Transactions of the Eleventh Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, Vol.A, 189-200, Academia: Prague.

MOUCHART, M., ROLIN, J.M. (1984) A note on conditional independence with statistical applications. *Statistica* **44** 557-584.

PEARL, J. (1988) *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufman: San Meteo, California.

SPEED, T.P. (1979) A note on nearest-neighbour Gibbs and Markov probabilities. *Sankkya Ser. A* **41** 184-197.

STUDENÝ, M. (1989) Multiinformation and the problem of characterization of conditional independence relations. *Problems of Control and Inf. Th.* **18** 3-16.

STUDENÝ, M. (1991) Conditional independence relations have no finite complete characterization. In *Transactions of the Eleventh Prague Conference on Information Theory, Statistical Decision Functions and Random Processes*, Vol.B, 377-396, Academia: Prague.

WHITTAKER, J. (1990) *Graphical Models in Applied Multivariate Statistics.* Wiley: Chichester.