3

Graphical Methods

Graphs whose nodes correspond to random variables are traditional tools for description of CI structures. One can distinguish three classic approaches: using *undirected graphs*, using *acyclic directed graphs* and using *chain graphs*. This chapter is an overview of graphical methods for describing CI structures with the main emphasis put on theoretical questions mentioned in Section 1.1. Both classic and advanced approaches are included. Note that elementary graphical concepts are introduced in Section A.3.

3.1 Undirected graphs

Graphical models based on undirected graphs are also known as *Markov networks* [100]. Given an undirected graph G over N one says that a disjoint triplet $\langle A,B|C\rangle \in \mathcal{T}(N)$ is *represented in* G, and writes $A \perp\!\!\!\perp B \,|\, C \; [G]$ if every route (equivalently every path) in G between a node in A and a node in B contains a node in C, that is, C separates between A and B in G. For illustration see Figure 3.1. Thus, every undirected graph G over N induces a formal independence model over N by means of the *separation criterion* (*for undirected graphs*):

$$\mathcal{M}_G = \{ \langle A,B|C\rangle \in \mathcal{T}(N) \,;\; A \perp\!\!\!\perp B \,|\, C \; [G] \} \,.$$

Let us call every independence model obtained in this way a *UG model*. These models were characterized by Pearl and Paz [99] in terms of a finite number of formal properties:

1. triviality $\quad A \perp\!\!\!\perp \emptyset \,|\, C \; [G]$,
2. symmetry $\quad A \perp\!\!\!\perp B \,|\, C \; [G]$ implies $B \perp\!\!\!\perp A \,|\, C \; [G]$,
3. decomposition $\quad A \perp\!\!\!\perp BD \,|\, C \; [G]$ implies $A \perp\!\!\!\perp D \,|\, C \; [G]$,
4. strong union $\quad A \perp\!\!\!\perp B \,|\, C \; [G]$ implies $A \perp\!\!\!\perp B \,|\, DC \; [G]$,
5. intersection $\quad A \perp\!\!\!\perp B \,|\, DC \; [G]$ and $A \perp\!\!\!\perp D \,|\, BC \; [G]$ implies $A \perp\!\!\!\perp BD \,|\, C \; [G]$,

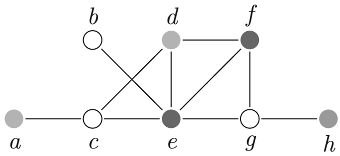

Fig. 3.1. The set $C = \{e, f\}$ separates between sets $A = \{a, d\}$ and $B = \{h\}$.

6. transitivity $\quad A \perp\!\!\!\perp B \,|\, C \; [G]$
implies $\; A \perp\!\!\!\perp \{d\} \,|\, C \; [G]$ or $\{d\} \perp\!\!\!\perp B \,|\, C \; [G]$.

This axiomatic characterization implies that every UG model is a graphoid satisfying the composition property.

Remark 3.1. Please note that the above-mentioned separation criterion was a result of some evolution. Theory of Markov fields stems from statistical physics [95] where undirected graphs were used to model geometric arrangements in space. Several types of Markov conditions were later introduced (see § 3.2.1 of Lauritzen [70] for an overview) in order to associate these graphs and probabilistic CI structures. The original "pairwise Markov property" was strengthened to the "local Markov property" and this was finally strengthened to the "global Markov property". These Markov conditions differ in general (e.g. [80]) but they coincide if we restrict our attention to positive measures [69].

The authors who contributed to the theory of Markov fields in the 1970s (see the introduction of Speed [120] for references) basically restricted their attention to the class of positive discrete probability measures. In other words, they used undirected graphs to describe structural properties of probability measures taken from this class; that is, they actually kept to a special *distribution framework* of positive discrete measures (see Section A.9.5 for an explanation of what I mean by a distribution framework). It was already found in the 1970s that the above-mentioned Markov conditions for undirected graphs are equivalent for the measures from the respective class Ψ of positive discrete measures over N. Moreover, it was shown later that the global Markov property, which is clearly the strongest one of those three Markov properties, cannot be strengthened within the framework of Ψ (see Remark 3.2 for more explanation).

Thus, the theory of Markov fields was developed under an implicit assumption that a particular distribution framework is considered. Undirected graphs also appeared in the 1970s in statistics (see Wermuth [156]) where they were used to describe so-called "covariance selection models" formerly introduced by Dempster [34]. However, statisticians restricted their attention to another class of probability measures, namely to the class of regular Gaussian measures (see p. 30). That means, they kept to another distribution framework. Nevertheless, it can be shown that the global Markov property is