Noname manuscript No.
(will be inserted by the editor)

Related definitions and notations

First Author · Second Author · Third Author

Received: date / Accepted: date

The graph theory definitions are adopted from [Diestel, 2000]. The Bayesian network definitions and notations are largely adopted from [Neapolitan, 2004].

1 Probability theory

- marginal probability distribution
- joint probability distribution
- conditional probability distribution
- Bayes' Theorem
- independent
- conditional independent
- prior propability
- posterior probability

2 Graphical model

- graph - nodes, vertices, variables are used interchangably; edges, arcs, links are used interchangably;
-

Definition 1 A **graph** is a pair $G = (V, E)$ comprising a set V of vertices (or nodes) together with a set E of edges (or arcs).

F. Author
first address
Tel.: +123-45-678910
Fax: +123-45-678910
E-mail: fauthor@example.com

S. Author
second address

The vertex set of a graph G is referred to as $V(G)$, its edge set as $E(G)$.

Definition 2 Two vertices x, y of G are **adjacent**, or **neighbours**, if xy is an edge of G.

Definition 3 Let $G = (V, E)$ and $G' = (V', E')$ be two graphs. If $V' \subseteq V$ and $E' \subseteq E$, then G' is a **subgraph** of G (and G is a **supergraph** of G'), written as $G' \subseteq G$. If $G' \subseteq G$ and G' contains all the edges $xy \in E$ with $x, y \in V'$, then G' is an **induced subgraph** of G, written as $G' = G[V']$.

Definition 4 A **path** in a graph G is a non-empty graph $P = (V, E)$ of the form

$$V = \{X_0, X_1, \dots, X_k\}, E = \{X_0X_1, X_1X_2, \dots, X_{k-1}X_k\},$$

where the X_is are all distinct. If $k \geq 3$ and $X_k = X_0$, then the graph $C = (V, E)$ is called a **cycle**.

Definition 5 A non-empty graph G is called **connected** if any two of its vertices are linked by a path in G.

Definition 6 An acyclic graph, one not containing any cycles, is called a **forest**. A connected forest is called a **tree**.

Definition 7 A **directed graph** (or **digraph**) is a pair (V, E), where V is a finite set of vertices, and E is a set of ordered pairs of distinct vertices in V.

Definition 8 A directed graph $G = (V, E)$ is called a directed acyclic graph (**DAG**) if it contains no directed cycles. Given nodes X_1 and X_2 in V, X_1 is called a **parent** of X_2 and X_2 is called a **child** of X_1 if there is a directed edge from X_1 to X_2. X_1 is called an **ancestor** of X_2 and X_2 is called a **descendent** of X_1 if there is a directed path from X_1 to X_2. X_2 is a **nondescendent** of X_1 if X_2 is not a descendent of X_1.

Definition 9 A **chain graph** [Frydenberg, 1990] is a graph which may have directed and undirected edges, but with no directed cycles.

Undirected graphs and DAGs are special cases of chain graphs. The definition was later extended to hybrid graphs with no directed or semi-directed cycles (i.e., no edges point the same direction).

Definition 10 A **hybrid graph (a.k.a., mixed graph)** is a graph consisting of directed and undirected edges.

Definition 11 A **collider** in a hybrid graph is a node with at least two parents.

Definition 12 Let $G_1 = (X, E_1)$ and $G_2 = (X, E_2)$ be two DAGs containing the same set of variables X. Then G_1 and G_2 are **Markov equivlent** if for every three mutually disjoint subsets $A, B, C \subseteq X$, A and B are d-separated by C in G_1 if and only if A and B are d-separated by C in G_2.

Proposition 1 *Let $G_1 = (X, E_1)$ and $G_2 = (X, E_2)$ be two DAGs containing the same set of variables X. Then the following are equivalent:*

1. *G_1 and G_2 are Markov equivalent.*
2. *Based on the Markov condition, they entail the same conditional independencies.*
3. *For every joint probability distribution P of X, $< G_1, P >$ satisfies the Markov condition if and only if $< G_2, P >$ satisfies the Markov condition.*
4. *They have the same skeleton and the same set of colliders.*

Definition 13 An edge is **compelled** in an equivalent class $\mathcal{G}$ if it exists in every DAG in $\mathcal{G}$.

Definition 14 An equivalence class $\mathcal{G}$ of directed acyclic graphs over X variables can be described by the **essential graph** [Andersson *et al.*, 1997] (or completed partially directed acyclic graph **CPDAG** [Chickering, 2002], **DAG pattern** [Neapolitan, 2004]) of $\mathcal{G}$ which is a hybrid graph G^* over X defined as follows:

- $X_i \to X_j \in E(G^*)$ iff $X_i \to X_j \in E(G)$ for every $G \in \mathcal{G}$,
- $X_i - X_j \in E(G^*)$ iff there are $G_1, G_2 \in \mathcal{G}$ such that $X_i \to X_j \in E(G_1)$ and $X_i \leftarrow X_j \in E(G_2)$.

This is a unique representation of an equivalent class.

Definition 15 The **moral graph** of a DAG is the skeleton of the hybrid graph obtained by adding an undirected edge between each pair of non-adjacent parents that have a common child.

References

[Andersson *et al.*, 1997] S. A. Andersson, D. Madigan, and M. D. Perlman. A characterization of Markov equivalence classes for acyclic digraphs. *The Annals of Statistics*, 25(2):505–541, 1997.

[Chickering, 2002] D. M. Chickering. Optimal structure identification with greedy search. *Journal of machine learning research*, 3(Nov):507–554, 2002.

[Diestel, 2000] R. Diestel. *Graph theory*. Springer-Verlag Berlin and Heidelberg GmbH & amp, 2000.

[Frydenberg, 1990] M. Frydenberg. The chain graph Markov property. *Scandinavian Journal of Statistics*, pages 333–353, 1990.

[Neapolitan, 2004] R. E. Neapolitan. *Learning bayesian networks*, volume 38. Pearson Prentice Hall Upper Saddle River, NJ, 2004.