Directed Graphical Models

Felipe José Bravo Márquez

August 13, 2021

Directed Graphical Models

- Probabilistic graphical models (PGMs) are a rich framework for encoding probability distributions over complex domains [Koller and Friedman, 2009].
- In this class we will focus on directed graphical models (DGMs), which are one type of PGM.
- Directed graphical models (DGMs) are a family of probability distributions that admit a compact parametrization that can be naturally described using a directed graph.
- DGMs are also known as Bayesian networks.
- Statistical inference for DGMs can be performed using frequentist or Bayesian methods, so it is misleading to call them Bayesian networks [Wasserman, 2013].

Directed Acyclic Graphs (DAGs)

- A directed graph consists of a set of nodes with arrows between some nodes.
- Graphs are useful for representing independence relations between variables.
- More formally, a directed graph G consists of a set of vertices V and an edge set E of ordered pairs of vertices.
- For our purposes, each vertex corresponds to a random variable.
- If $(Y, X) \in E$ then there is an arrow pointing from Y to X.

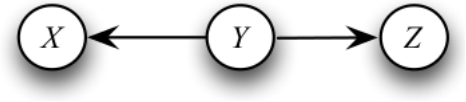


Figure: A directed graph with vertices $V = \{X, Y, Z\}$ and edges $E = \{(Y, X), (Y, Z)\}$.

Directed Acyclic Graphs (DAGs)

- If an arrow connects two variables X and Y (in either direction) we say that X and Y are adjacent.
- If there is an arrow from X to Y then X is a parent of Y and Y is a child of X.
- The set of all parents of X is denoted by π_X or $\pi(X)$.
- A directed path between two variables is a set of arrows all pointing in the same direction linking one variable to the other such as the chain shown below:

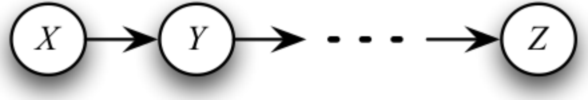
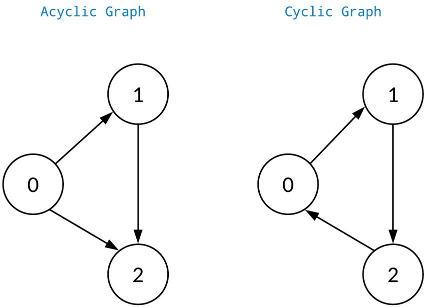


Figure: A chain graph with a directed path.

- X is an ancestor of Y if there is a directed path from X to Y (or X = Y).
- We also say that Y is a descendant of X.

Directed Acyclic Graphs (DAGs)

- A directed path that starts and ends at the same variable is called a cycle.
- A directed graph is acyclic if it has no cycles.
- In this case we say that the graph is a directed acyclic graph or DAG.



- From now on, we only deal with directed acyclic graphs since it is very difficult to provide a coherent probability semantics over graphs with directed cycles.

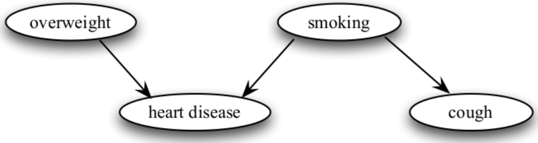
Probability and DAGs

- Let G be a DAG with vertices $V = (X_1, \ldots, X_d)$.
- If P is a distribution for V with probability function $f(x)$ (density or masss), we say that G represents P , if

$$f(x) = \prod_{j=1}^{d} f(x_j|\pi_{x_j})$$

where π_{x_j} is the set of parent nodes of X_j

- The next figute shows a DAG with four variables.



- The probability function takes the following decomposition:
- f(overweight, smoking, heart disease, cough) $=$ f(overweight) $\times$ f(smoking) $\times$ f(heart, disease$|$ overweight, smoking) $\times$ f(cough$|$smoking).

Conditional Independence

- Let X, Y and Z be random variables.
- X and Y are conditionally independent given Z, written $X \perp Y | Z$, if:

$$f(x, y|z) = f(x|z)f(y|z)$$

  for all x, y and z.
- Notice that f can be either a density function for continous random variables or a probability mass function for discrete random variables.
- Intuitively, this means that, once you know Z, Y provides no extra information about X.

D-separation

- sdsad

Conclusions

- Blabla

References I

Koller, D. and Friedman, N. (2009).
Probabilistic graphical models: principles and techniques.
MIT press.

Wasserman, L. (2013).
All of statistics: a concise course in statistical inference.
Springer Science & Business Media.