Graphical models

Introduction to Machine Learning (CSCI 1950-F), Summer 2011

Solve the following problems. Provide mathematical justification for your answers.

Recall that if $X_1, \ldots, X_n$ have joint distribution p (where p is a PMF or PDF), and G is an ordered DAG (directed acyclic graph) with n vertices, we say that p *respects* G if

$$p(x_1, \ldots, x_n) = \prod_{i=1}^{n} p(x_i | x_{pa(i)})$$

for all $x_1, \ldots, x_n$. (Here, $pa(i)$ denotes the parents of vertex i in G, and $x_{pa(i)}$ denotes the tuple $(x_j : j \in pa(i))$. Also, we say that a DAG G is *ordered* if its vertices are numbered $1, \ldots, n$ such that if i is a parent of j then $i < j$.)

Note: For each of the following two exercises, a formal argument would require a proof by induction. However, it is acceptable for you to give an informal argument if you are not familiar with proof by induction (or if you prefer to give an informal argument). If you wish, you can prove these results just in the case of $n = 4$.

(1) Normalization

Suppose that you have an ordered DAG with n vertices and you want to construct a joint distribution for discrete random variables $X_1, \ldots, X_n$ by specifying each of the conditional PMFs $p(x_i | x_{pa(i)})$. Suppose that for each $i = 1, \ldots, n$, $p(x_i | x_{pa(i)})$ is a PMF (i.e. it is nonnegative and $\sum_{x_i} p(x_i | x_{pa(i)}) = 1$). Show that the function g defined by

$$g(x_1, \ldots, x_n) = \prod_{i=1}^{n} p(x_i | x_{pa(i)})$$

is also a PMF, by showing that

$$\sum_{x_1} \cdots \sum_{x_n} g(x_1, \ldots, x_n) = 1.$$

(Hint: You will have to think about the graph.)

(2) Complete DAGs

A *complete directed graph* is a directed graph in which every pair of distinct vertices is connected by exactly one edge. A *complete DAG* is a complete directed graph that is acyclic. Show that any joint distribution for random variables $X_1, \ldots, X_n$ respects any ordered complete DAG G with n vertices. (Hint: Use the definition of an "ordered" DAG.)