Homework V: Stat 246

Due, Thursday, May 15

1. (60 points) Imagine a graph over random *binary* variables $X_i, i = 0, \ldots, 2$, $X_i \in \{0, 1\}$, where edges are between consecutive nodes, i.e.

$$V = \{0, 1, 2\}, \quad E = \{(0, 1), (1, 2)\}.$$

The cliques in this graph are singletons and consecutive pairs.

   (a) If (i, j) are a clique let $\psi_{ij}(X_i, X_j) = \alpha X_i X_j$, for some $\alpha > 0$. Compute the normalizing constant Z. Compute the marginal distribution on X_0.

   (b) What happens to the joint distribution as $\alpha \to 0$.

   What happens to the joint distribution as $\alpha \to \infty$.

   (c) Define the matrix $M = \begin{pmatrix} 1 & 1 \\ 1 & e^{\alpha} \end{pmatrix}$. Find an expression for Z in terms of M.

   (d) Imagine now the same type of graph on n nodes. Only consecutive nodes are connected with an edge, with the same clique function as above. Generalize the formula for Z in terms of M.

   Can you derive a formula for the marginal distribution on X_0.

   (e) For any joint distribution $p(x_1, \ldots, x_n)$ on n variables show that

$$p(x_1, \ldots, x_n) = p(x_n | x_1, \ldots, x_{n-1}) p(x_{n-1} | x_1, \ldots, x_{n-2}) \cdots p(x_2 | x_1) p(x_1).$$

   Use this formula and the conditional independence properties defined by the graph above to show that the joint distribution defines a Markov Chain.

   Is it homogeneous, i.e. are the transition probabilities the same at each step?

   (f) Now assume the same graphical model and the same clique function but the variables $X_i, i = 0, 1, 2$ are continuous and supported on the entire real line can you compute a normalizing constant?

   If you add a clique function for each singleton defined as $\psi_i(X_i) = -\beta X_i^2$ for $\beta > \alpha/2$. Can you compute a normalizing constant now? What is the joint distribution.

2. Consider the graph $G = (V, E)$,

$$V = \{1, 2, 3, 4, 5, 6\}, \; E = \{(1, 2), (2, 3), (3, 4), (3, 5), (5, 6)\}.$$

   (a) Draw the graph.

(b) Let $X_i, i = 1, \ldots, 6$ be discrete random variables with state space $\{1, \ldots, K\}$, whose joint distribution $p(X_1, \ldots, X_6)$ is Markov with respect to the graph G. Write the general formula for p in terms of the cliques of G.

(c) How many operations would it take to compute the normalizing constant Z by brute force summation over all sequences $(x_1, \ldots, x_6)$?

(d) Describe in detail a faster method that would only take order $6K^2$ operations.