UC Berkeley
Department of Electrical Engineering and Computer Science
Department of Statistics

EECS 281A / STAT 241A Statistical Learning Theory

Problem Set 3
Fall 2012

Issued: Tues. Sep. 18, 2012 **Due:** Tues. Oct 2, 2012

Reading: Chapters 2, 3 and 4

Problem 3.1
Suppose that three discrete random variables (X, Y, Z) have a joint PMF such that $p(x, y, z) > 0$ for all (x, y, z). Show that if $X \perp Y \mid Z$ and $X \perp Z \mid Y$, then we have $X \perp (Y, Z)$. Is this still true if we allow $p(x, y, z) = 0$ for some (x, y, z)?

Problem 3.2
For each of the following statements, either give a proof of its correctness, or a counterexample to show incorrectness.

(a) If $X_1 \perp X_2$, then $X_1 \perp X_2 \mid X_3$.

(b) If $X_1 \perp X_2 \mid X_4$ and $X_1 \perp X_3 \mid X_4$, then $X_1 \perp (X_2, X_3) \mid X_4$.

(c) If $X_1 \perp (X_2, X_3) \mid X_4$, then $X_1 \perp X_2 \mid X_4$.

Problem 3.3
Graphs and independence relations: For $i = 1, 2, 3$, let X_i be an indicator variable for the event that a coin toss comes up heads (which occurs with probability q). Supposing that that the X_i are independent, define $Z_4 = X_1 \oplus X_2$ and $Z_5 = X_2 \oplus X_3$ where $\oplus$ denotes addition in modulo two arithmetic.

(a) Compute the conditional distributions of (X_2, X_3) given $Z_5 = 0$ and $Z_5 = 1$ respectively.

(b) Draw a directed graphical model (the graph and conditional probability tables) for these five random variables. What independence relations does the graph imply?

(c) Draw an undirected graphical model (the graph and compatibility functions) for these five variables. What independence relations does it imply?

(d) Under what conditions on q do we have $Z_5 \perp X_3$ and $Z_4 \perp X_1$? Are either of these marginal independence assertions implied by the graphs in (b) or (c)?

Problem 3.4
Consider a sequence of random variables $(X_1, \ldots, X_d)$ generated according to the following procedure:

(i) Sample $X_1 \sim N(0, 1)$.

(ii) Given some $a \in (-1, 1)$, for $t = 1, \ldots, d-1$, set $X_{t+1} = aX_t + \sqrt{1-a^2}\, W_t$, where the $\{W_t\}_{t=1}^{d-1}$ are independent $N(0, 1)$ variables, with W_t chosen independently of X_t.

(a) Compute the covariance matrix $\Sigma \in \mathbb{R}^{d \times d}$ of the random vector $X \in \mathbb{R}^d$.

(b) Show that the inverse covariance matrix Σ^{-1} is always tridiagonal, meaning that it is non-zero only on its diagonal and on the two diagonals above and below the main diagonal. (I.e., $(\Sigma^{-1})_{ij} = 0$ for all $|i - j| > 1$.)

(*Hint:* You may want to simulate this numerically just to confirm the intuition. In proving the result, the Hammersley-Clifford theorem could be helpful.)

Problem 3.5
Consider the directed graph shown in Figure 1(a). For each of the following conditional independence statements, verify whether or not they hold. In each case, be explicit using the Bayes ball algorithm, indicating how the ball gets through, or how it is blocked for each possible path.

(a) $X_2 \perp X_8 \mid \{X_3, X_4, X_5\}$.

(b) $X_8 \perp X_9 \mid \{X_3, X_4, X_5\}$.

(c) $X_7 \perp X_{10} \mid \{X_3, X_4, X_5\}$.

Problem 3.6
Undirected graphs and elimination: Consider the undirected graph in Figure 1(b): it is a 3×3 grid or lattice graph.

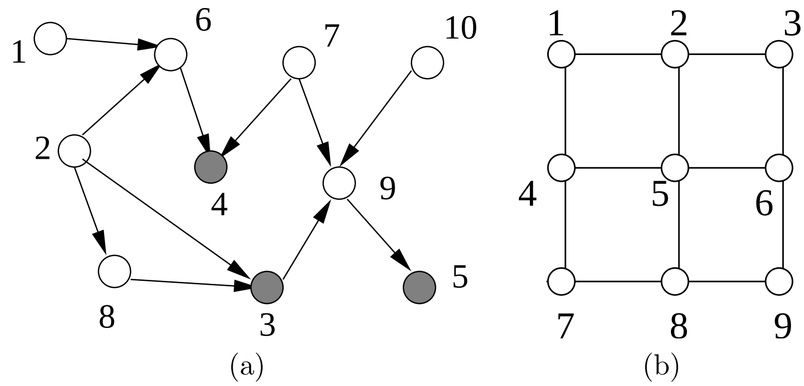

Figure 1: (a) A directed graph. (b) An undirected graphical model: a 3×3 grid, frequently used in spatial statistics and image processing.

(a) Sketch the sequence of graphs obtained by running the algorithm `Graph-eliminate`:

(i) Following the ordering $\{5, 4, 8, 6, 2, 9, 3, 7, 1\}$?

(ii) Following the ordering $\{1, 7, 3, 9, 2, 4, 6, 8, 5\}$?

What is the largest clique formed by each graph sequence? Which ordering is preferable?

(b) Using intuition from the previous example ($n = 3$), give a reasonable ($\ll n^2$) upper bound on the treewidth of the $n \times n$ grid.