6.892 Machine learning seminar

Prof. Tommi S. Jaakkola
tommi@ai.mit.edu

February 16, 1999

A simple node elimination algorithm

(Find anything wrong with this handout?)

Assume we have an undirected graph $G^{(n)}$ consisting of n nodes. We denote the set of cliques in this graph by $C^{(n)} = \{c_1, \ldots, c_k\}$. Each c_k is thus a (maximal) complete subset of nodes. The most general form of the joint distribution that agrees with this graph is given by

$$P(x_1, \ldots, x_n) = \prod_{c \in C^{(n)}} \phi_c^{(n)}(x_c) \tag{1}$$

where $x_c = \{x_i\}_{i \in c}$, the set of variables associated with the nodes in the clique.

We define now a recursive node elimination algorithm for computing marginal probabilities (e.g. $P(x_1)$) from this joint distribution. The algorithm proceeds by eliminating all the nodes in the graph except those that we are interested in (e.g. x_1).

1. Identify a node in $G^{(n)}$ to be eliminated, say j.
2. Find all the cliques that contain this node and call this set of cliques $C_j^{(n)}$.
3. Let $\tilde{c}$ be the set of neighbors of j (note: doesn't include j itself).
4. Construct a potential function over the set $\tilde{c}$:

$$\phi_{\tilde{c}}^{(n)}(x_{\tilde{c}}) = \sum_{x_j} \prod_{c \in C_j^{(n)}} \phi_c(x_c) \tag{2}$$

5. Create a new graph $G^{(n-1)}$ by first connecting all the neighbors of j in $G^{(n)}$ and then removing the node j and all its associated edges.

6. Find the set of cliques $C^{(n-1)}$ of the new graph $G^{(n-1)}$.

7. Assign the potential functions $\phi_c^{(n)}(x_c)$, $c \in \{C^{(n)} \setminus C_j^{(n)}, \tilde{c}\}$, to the new cliques in $C^{(n-1)}$ such that the variables in these functions are contained in the cliques they are assigned to. Note: each potential function goes to only one of the new cliques.

8. Create the potential functions $\phi^{(n-1)}$ for the new graph by multiplying the functions $\phi_c^{(n)}(x_c)$ that fell in the same clique. In other words, for each $c \in C^{(n-1)}$

$$\phi_c^{(n-1)}(x_c) = \prod_{c'} \phi_{c'}^{(n)}(x_{c'}) \tag{3}$$

   where the product is over those $c' \in \{C^{(n)} \setminus C_j^{(n)}, \tilde{c}\}$ that were assigned to c.

   The joint distribution corresponding to $G^{(n-1)}$ is now

$$P(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) = \prod_{c \in C^{(n-1)}} \phi_c^{(n-1)}(x_c) \tag{4}$$

9. Go to step 1 until we are left with only the variables of interest.

Choice of ordering:

- The choice of ordering can affect the complexity of the marginalization considerably
- Simple heuristic: choose always the node with the fewest neighbors in the remaining graph, breaking ties arbitrarily. (This assumes that the number of values that each variable can take is the same throughout the graph).