Advanced Methods in Machine Learning - HW2

Amir Bar 301804191, Ben Bogin 200375152

April 29, 2017

Question 2

(a) We want to show:

$$\mu_{ij}(x_i, x_j) \propto \phi_{ij}(x_i, x_j) \prod_{k \in N(i)/j} m_{ki}(x_i) \prod_{k \in N(j)/i} m_{kj}(x_j)$$

Proof:

Define T_i to be all nodes "before" i and E_i the edges in this subtree and likewise for T_j, E_j.

$$\begin{aligned} p_{ij}(x_i, x_j) &= \sum_{x_1,\ldots,x_n/x_i,x_j} p(X_i = x_i, X_j = x_j, x_1, \ldots, x_n/x_i, x_j) \\ &\propto \sum_{x_1,\ldots,x_n/x_i,x_j} \left(\prod_{kl \in E/ij} \phi_{kl}(x_k, x_l) \right) \cdot \phi_{ij}(x_i, x_j) \\ &= \phi_{ij}(x_i, x_j) \left(\sum_{T_i} \prod_{kl \in E_i} \phi_{kl}(x_k, x_l) \right) \left(\sum_{T_j} \prod_{kl \in E_j} \phi_{kl}(x_k, x_l) \right) \end{aligned}$$

Where the last equation above is legal since this is a tree, so we can separate the independent nodes before i and the nodes after j.

We can further separate $(\sum \prod \phi_{kl}(x_k, x_l))$ to each one of $k \in N(i)$:

$$\prod_{k \in N(i)} \left(\sum_{T_k} \prod_{kl \in E_i} \phi_{ki}(x_k, x_i) \right) = \prod_{k \in N(i)} m_{ki}(x_i)$$

And then we get:

$$p_{ij}(x_i, x_j) \propto \phi_{ij}(x_i, x_j) \prod_{k \in N(i)/j} m_{ki}(x_i) \prod_{k \in N(j)/i} m_{kj}(x_j)$$

(b) We want to show:

$$p(x_1, \ldots, x_n) \propto \prod_i \mu_i(x_i) \prod_{ij} \frac{\mu_{ij}(x_i, x_j)}{\mu_i(x_i)\mu_j(x_j)}$$

Proof:

$$
\begin{aligned}
&\prod_i \mu_i(x_i) \prod_{ij} \frac{\mu_{ij}(x_i, x_j)}{\mu_i(x_i)\mu_j(x_j)} \\
&\propto \left(\prod_i \prod_{k \in N(i)} m_{ki}(x_i) \right) \left(\prod_{ij} \frac{\phi_{ij}(x_i, x_j) \prod_{k \in N(i)/j} m_{ki}(x_i) \prod_{k \in N(j)/i} m_{kj}(x_j)}{\prod_{k \in N(i)} m_{ki}(x_i) \prod_{k \in N(j)} m_{kj}(x_j)} \right) \\
&= \left(\prod_i \prod_{k \in N(i)} m_{ki}(x_i) \right) \left(\prod_{ij} \frac{\phi_{ij}(x_i, x_j)}{m_{ji}(x_i) m_{ij}(x_j)} \right)
\end{aligned}
$$

It holds that

$$\prod_i \prod_{k \in N(i)} m_{ki}(x_i) = \prod_{ij} m_{ji}(x_i) m_{ij}(x_j)$$

Therefore,

$$\prod_i \mu_i(x_i) \prod_{ij} \frac{\mu_{ij}(x_i, x_j)}{\mu_i(x_i)\mu_j(x_j)} \propto \prod_{ij} \phi_{ij}(x_i, x_j) \propto p(x_1, ..., x_n)$$

Thus we get

$$p(x_1, ..., x_n) \propto \prod_i \mu_i(x_i) \prod_{ij} \frac{\mu_{ij}(x_i, x_j)}{\mu_i(x_i)\mu_j(x_j)}$$