Factorization of KL-divergence for tree-structured distributions

Charles Sutton

20 July 2005

If two distributions $p(\mathbf{x})$ and $q(\mathbf{x})$ factorize according to the same tree-structured graph $G = (V, E)$, then their KL-divergence can be written in terms of divergences between pairwise marginals. To see this, observe that the sequence distributions can be written

$$p(\mathbf{x}) = \frac{\prod_{(s,t)\in E} p(x_s, x_t)}{\prod_{s\in V} p(x_s)^{d_s - 1}}, \tag{1}$$

and similarly for q. This follows from the junction tree theorem; consider the junction tree whose nodes corresponds to edges of G, and whose sepsets are single vertices.

Then we can factorize the KL-divergence as follows:

$$\mathrm{KL}(q(\mathbf{x}) \| p(\mathbf{x})) = \sum_{\mathbf{x}} q(\mathbf{x}) \log q(\mathbf{x}) - \sum_{\mathbf{x}} q(\mathbf{x}) \log p(\mathbf{x}) \tag{2}$$

$$= \sum_{\mathbf{x}} q(\mathbf{x}) \log \left(\frac{\prod_{(s,t)\in E} q(x_s, x_t)}{\prod_{s\in V} q(x_s)^{d_s - 1}} \right) - \sum_{\mathbf{x}} q(\mathbf{x}) \log \left(\frac{\prod_{(s,t)\in E} p(x_s, x_t)}{\prod_{s\in V} p(x_s)^{d_s - 1}} \right) \tag{3}$$

$$\begin{aligned} = &\sum_{(s,t)\in E} \sum_{\mathbf{x}} q(\mathbf{x}) \log q(x_s, x_t) - \sum_{s\in V} (d_s - 1) \sum_{\mathbf{x}} q(\mathbf{x}) \log q(x_s) \\ &- \sum_{(s,t)\in E} \sum_{\mathbf{x}} q(\mathbf{x}) \log p(x_s, x_t) + \sum_{s\in V} (d_s - 1) \sum_{\mathbf{x}} q(\mathbf{x}) \log p(x_s) \end{aligned} \tag{4}$$

$$\begin{aligned} = &\sum_{(s,t)\in E} \sum_{x_s, x_t} q(x_s, x_t) \log q(x_s, x_t) - \sum_{s\in V} (d_s - 1) \sum_{x_s} q(x_s) \log q(x_s) \\ &- \sum_{(s,t)\in E} \sum_{x_s, x_t} q(x_s, x_t) \log p(x_s, x_t) + \sum_{s\in V} (d_s - 1) \sum_{x_s} q(x_s) \log p(x_s) \end{aligned} \tag{5}$$

$$= \sum_{(s,t)\in E} \mathrm{KL}(q(x_s, x_t) \| p(x_s, x_t)) - \sum_{s\in V} (d_s - 1) \mathrm{KL}(q(x_s) \| p(x_s)), \tag{6}$$

which is the sort of equation I promised. Equation 5 follows from the standard trick that $\sum_{\mathbf{x}} q(\mathbf{x}) \log q(x_s) = \sum_{x_s} \log q(x_s) \sum_{\mathbf{x}\setminus x_s} q(\mathbf{x}) = \sum_{x_s} q(x_s) \log q(x_s)$. If you want to calculate the KL divergence exactly in a loopy graph, then apply this method to the junction tree.