Homework #1

Due: Th. Jan.22

Learning in Graphical Models ICS 280
URL: http://www.ics.uci.edu/~welling/teaching/GraphicalModels.html
Instructor :Max Welling
Office: 414C
phone: 824-8169
email: welling@ics.uci.edu

1. Read the handouts provided in class and the classnotes and book-chapters provided on the class web-page.
2. a) Consider the following model with input variable x, output (or target) variable t and hidden (or unobserved) variable y, and three parameters $\theta_1, \theta_2, \theta_3$,

$$p(t, y, x) = p_{\theta_1}(t|y)\; p_{\theta_2}(y|x)\; p_{\theta_3}(x) \tag{1}$$

   Use Bayes' rule to compute the distribution,

$$p(y|x, t) \tag{2}$$

   where the final expression must be in terms of the model distributions $p_{\theta_1}, p_{\theta_2}, p_{\theta_3}$ only.

   b) Draw the graphical model and check whether x and t are independent, given y. Check this by explicit calculation, i.e. check if $p(t|y, x) = p(t|y)$.

   c) We now receive an IID data-set $\{t_i, x_i\}, \;\; i = 1..N$ (i.e. all data cases are drawn independently from the same distribution). Write the expression for the maximum likelihood objective function in terms of $p_{\theta_1}, p_{\theta_2}, p_{\theta_3}$ and data $\{t_i, x_i\}, \;\; i = 1..N$. Note: we did not receive data for y, so the final expression cannot depend on it. (hint: use marginalization).

   b) A Bayesian come along and provides you with priors for the parameters: $p(\theta_1), p(\theta_2), p(\theta_3)$. The Bayesian expression for the probability distribution becomes,

$$p(t, y, x, \theta_1, \theta_2, \theta_3) = p(t|y, \theta_1)\; p(y|x, \theta_2)\; p(x|\theta_3)\; p(\theta_1)\; p(\theta_2)\; p(\theta_3) \tag{3}$$

   Draw the corresponding graphical model. Image we observe x and t, so we shade these nodes. Now check if the following statements are true,

$$\theta_1 \perp \theta_2 | \{t, x\} \tag{4}$$

$$\theta_1 \perp \theta_3 | \{t, x\} \tag{5}$$

Bonus: Use Bayes' rule to express the posterior distribution,

$$p(\theta_1, \theta_2, \theta_3 | \{x_i, t_i\}) \tag{6}$$

in terms of the model distributions $p(t|y, \theta_1)$, $p(y|x, \theta_2)$, $p(x|\theta_3)$, $p(\theta_1)$, $p(\theta_2)$, $p(\theta_3)$ and the data $\{t_i, x_i\}, \;\; i = 1..N$.

3. a) Draw the directed graph corresponding to the following expression:

$$p(x_1, x_2, x_3, x_4) = p(x_4|x_2, x_3)\, p(x_2|x_1)\, p(x_3|x_1)\, p(x_1) \tag{7}$$

b) Which of the following statements are true:
(i) $x_1 \perp x_4 | \{x_2, x_3\}$
(ii)$x_2 \perp x_3 | \{x_1, x_4\}$

c) Remove the arrows from graph, converting it into an undirected graph. Answer item b) again for the undirected graph.

d) Write the most general expression for the probability distribution as a product of factors $\psi_A(x_A)$, where A denotes a subset of the random variables in the graph. This expression should be consistent with all conditional independencies implied by the undirected graph.

e) In both the directed and the undirected graph, add an edge between variable x_1 and x_4. In the directed case, can we direct the edge in both directions (explain).

f) In both cases, consider once more the conditional independence statements from b).