CSC412 / CSC2506 Sample Problems for Midterm

1. Recall that the definition of an exponential family model is:

$$f(x|\eta) = h(x)g(\eta)\exp(\eta^\top T(x))$$

where:

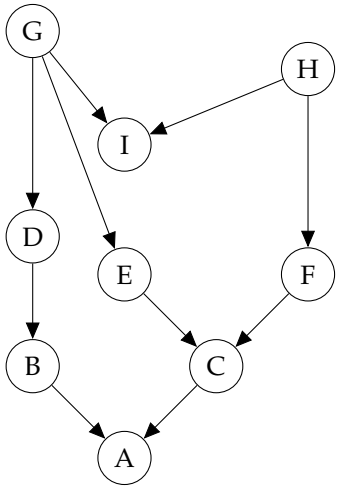
η	are the parameters
$T(x)$	are the sufficient statistics
$h(x)$	is the base measure
$g(\eta)$	is the normalizing constant

Consider the univariate Gaussian, with mean μ and precision $\lambda = \frac{1}{\sigma^2}$:

$$p(D|\mu,\lambda) = \prod_{i=1}^{N}(\frac{\lambda}{2\pi})^{\frac{1}{2}}\exp(-\frac{\lambda}{2}(x_i-\mu)^2)$$

What are η and $T(x)$ for this distribution when represented in exponential family form?

2. Consider the following directed graphical model:

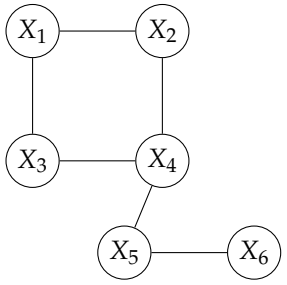


(a) List all variables that are independent of A given evidence on B.

(b) Write down the factorized normalized joint distribution that this graphical model represents.

(c) If each node is a single discrete random variable in $\{1, ..., K\}$ how many distinct joint states can the model take? That is, how many different configurations can the variables in this model be set?

3. (Murphy 20.1) Variable Elimination



a. Suppose we want to compute the partition function using the elimination ordering $\prec = (1, 2, 3, 4, 5, 6)$, i.e.,

$$\sum_{x_6}\sum_{x_5}\sum_{x_4}\sum_{x_3}\sum_{x_2}\sum_{x_1}\psi_{12}(x_1,x_2)\psi_{13}(x_1,x_3)\psi_{24}(x_2,x_4)\psi_{34}(x_3,x_4)\psi_{45}(x_4,x_5)\psi_{56}(x_5,x_6) \quad (20.71)$$

If we use the variable elimination algorithm, we will create new intermediate factors. What is the largest intermediate factor?

b. Add an edge to the original MRF between every pair of variables that end up in the same factor. (These are called fill in edges.) Draw the resulting MRF. What is the size of the largest maximal clique in this graph?

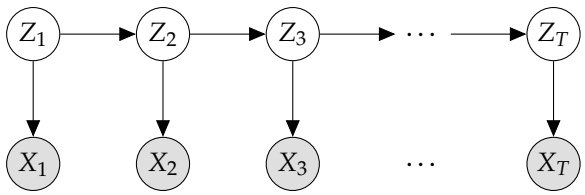
c. Now consider elimination ordering $\prec = (4, 1, 2, 3, 5, 6)$, i.e.,

$$\sum_{x_6}\sum_{x_5}\sum_{x_3}\sum_{x_2}\sum_{x_1}\sum_{x_4}\psi_{12}(x_1,x_2)\psi_{13}(x_1,x_3)\psi_{24}(x_2,x_4)\psi_{34}(x_3,x_4)\psi_{45}(x_4,x_5)\psi_{56}(x_5,x_6) \quad (20.72)$$

If we use the variable elimination algorithm, we will create new intermediate factors. What is the largest intermediate factor?

d. Add an edge to the original MRF between every pair of variables that end up in the same factor. (These are called fill in edges.) Draw the resulting MRF. What is the size of the largest maximal clique in this graph?

4. Consider the Hidden Markov Model



(a) Assume you are able to sample from these conditional distributions, i.e.

$$x_i \sim p(X_i|\text{parents of } X_i).$$

Write down a step-by-step process to produce a sample observation from this model, i.e. $(x_1, x_2, x_3, \ldots, x_T)$ in terms of samples from the individual factors.