CS60073: Advanced Machine Learning

End-Autumn Semester Exam

Time: 3:00hrs　　　Answer all FOUR questions.　　　Max mark: $25 \times 4 = 100$

1.A. Consider the Markov chain with three states, $S = \{1, 2, 3\}$, that has the following transition matrix.

$$T = \begin{bmatrix} 1/2 & 1/4 & 1/4 \\ 1/3 & 0 & 2/3 \\ 1/2 & 1/2 & 0 \end{bmatrix}$$

Draw the state transition diagram for this chain. Let X_t denote the state at time t. If we know $P(X_1 = 1) = P(X_1 = 2) = 1/4$, find $P(X_1 = 3, X_2 = 2, X_3 = 1)$. [10]

1.B. We would like to sample from a distribution $p(x, y) = \exp(-x^2 - (y - x^2)^2)$, where $x, y \in \mathbb{R}$. We have access to a Gaussian and an uniform random number generator. [15]

(i) Suggest a proposal distribution and derive the acceptance probability for the Metropolis-Hastings sampler that might be used for this purpose.

(ii) Write down the steps of Gibbs sampling of $p(x, y)$. Derive the conditional probability $p(y|x)$ to be used in the Gibbs sampler.

(iii) Describe details of a rejection sampling technique for $p(y|x)$ in the Gibbs sampler.

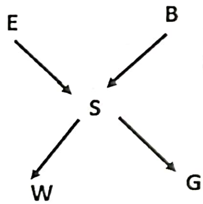
2.A. Define the Kullback-Leibler divergence between two distributions p and q. [5]

2.B. Pose variational inference on distributions p and q, involving observed variable X and latent variable Z, as an optimization problem. [5]

2.C. Show that the evidence lower bound ELBO $\mathcal{L}(q(Z)) = \int q(Z) \log \frac{p(X,Z)}{q(Z)} dZ$ is indeed a lower bound of the evidence $\log p(X)$. What is the modified optimization problem obtained due to this result? [10]

2.D. Clearly state the assumptions made in mean field variational inference. [5]

3. Consider the Bayesian network of five binary variables E, B, S, W, G given below.



E	e	-e
	1/10	9/10

B	b	-b
	1/10	9/10

S	s	-s
e ∧ b	9/10	1/10
e ∧ -b	2/10	8/10
-e ∧ b	8/10	2/10
-e ∧ -b	0	1

W	w	-w
s	8/10	2/10
-s	2/10	8/10

G	g	-g
s	1/2	1/2
-s	0	1

A. Compute the following probabilities: (show your calculation) [10]

(i) Given that G is true, what is the probability that S is true?

(ii) If B is true and E is false, what is the probability that W is true?

B. Does the message passing belief propagation algorithm perform exact inference for this Bayesian network? Explain your answer. [5]

C. We would like to perform approximate inference by Gibbs sampling for this network. A topologically sorted variable ordering is used. The variable E is observed to be true. What is the probability that - [10]

(i) The first sample is: E, B, W, G are true, and S is false.

(ii) The second sample is: E, B, G are true, and S, W are false.

[P.T.O]

4.A. What is an independence map (I-Map) for a given distribution p? [5]

4.B. Draw a Bayesian network which is an I-Map for the distribution $P(A, \ldots, J) = P(A|C, D, H)P(B|D, E, G)P(C|F, I)P(D|G, H)P(E|J)P(F|I)P(G|H, J)P(H)P(I)P(J)$. Check if the following variable pairs are d-separated: (i) E and I, (ii) B and H, and (iii) C and G. [10]

4.C. Draw a factor graph representation of the above Bayesian network. [5]

4.D. If possible, draw an undirected Markov Random Field to represent the above distribution. Define appropriate clique potentials. [5]

— BEST WISHES —