Probabilistic Graphical Models

HWK#3 (Part B)
Assigned Sunday, 8, 2020
Due: Tuesday, April 7, 2020

Problem 7 Variational Inference (15 points)

The file p.mat contains a distribution $p(x, y, z)$ on ternary state variables. Using BRML-toolbox, find the best approximation $q(x, y)q(z)$ that minimizes the Kullback-Leibler divergence $KL(q|p)$ and state the value of the minimal Kullback-Leibler divergence for the optimal q.

Problem 8 Variational Inference (10 points)

Consider the pairwise Markov network defined on a 2×2 lattice, as given in the file pMRF.mat.

1. Find the optimal fully factorized approximation $\prod_{i=1}^{4} q_i^{BP}$ by loopy belief propagation, based on the factor graph formalism.

2. Find the optimal fully factorized approximation $\prod_{i=1}^{4} q_i^{MF}$ by solving the variational mean-field equations.

3. By pure enumeration, compute the exact marginals p_i.

4. Averaged over all 4 variables, compute the mean expected deviation in the marginals

$$\frac{1}{4}\sum_{i=1}^{4}\frac{1}{2}\sum_{j=1}^{2}|q_i(x=j) - p_i(x=j)|$$

   for both the BP and MF approximations, and comment on your results.

Problem 9 Variational Inference (14 points)

Note that in this question, the notation $< \cdot >_q$ denotes averaging with respect to distribution q. *Consider the average of a positive function $f(x)$ with respect to a distribution $p(x)$:*

$$J = \log \int_x p(x) f(x)$$

where $f(x) \geq 0$. The simplest version of Jensen's inequality states that

$$J \geq \int_x p(x) \log f(x)$$

1. *By considering a distribution $r(x) \propto p(x) f(x)$, and $KL(q|r)$, for some variational distribution $q(x)$, show that*

$$J \geq -KL(q(x)|p(x)) + \langle \log f(x) \rangle_{q(x)}$$

The bound saturates when $q(x) \propto p(x) f(x)$. This shows that if we wish to approximate the average J, the optimal choice for the approximating distribution $q(x)$ depends on both the distribution $p(x)$ and integrand $f(x)$.

2. *Furthermore, show that*

$$J \geq -KL(q(x)|p(x)) - KL(q(x)|f(x)) - H(q(x))$$

where $H(q(x))$ is the entropy of $q(x)$. The first term encourages q to be close to p. The second encourages q to be close to f, and the third encourages q to be sharply peaked.