Takehome Exam
Topics in Optimization
(22-672)

Date: 1400.04.19

Time: Start (9am 1400.04.19) – Stop (11:00pm 1400.04.22)

- ***You may consult anyone or anything except your classmates!**
- ***You ought to e-mail your answer-sheets as PDF files to me through** `daneshgar@sharif.ir`. **Make sure that the name of the file is your student number and the subject is "TiO Takehome".**
- ***Unreadable documents will not go through the grading procedure.**
- ***You must explicitly state any result you may be using with appropriate references and attaching the necessary documents.**

1. Consider the factor graph of Figure 1 representing the probability distribution

$$p(x_1, x_2, z_1, z_2) = p_0(z_1)p_1(x_1, z_1)p_2(z_1, z_2)p_3(x_2, z_2) = e^{-H(x_1,x_2,z_1,z_2)},$$

in which

$$-H(x_1, x_2, z_1, z_2) \stackrel{\text{def}}{=} \log(p_0(z_1))+\log(p_1(x_1, z_1))+\log(p_2(z_1, z_2))+\log(p_3(x_2, z_2)).$$

In what follows, for Parts (a) and (b) assume that z_i is the output of a Markov chain on two states *Hot* and *Cold* with the following transition probabilities,

$$p_M(Hot|Hot) = 0.6, \quad p_M(Cold|Hot) = 0.4,$$

$$p_M(Cold|Cold) = 0.5, \quad \text{and} \quad p_M(Hot|Cold) = 0.5,$$

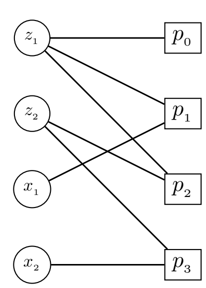

Figure 1: The factor graph.

and the initial distribution,

$$p_0(Hot) = 0.8 \quad \text{and} \quad p_0(Cold) = 0.2.$$

Also, assume that

$$p_1(x_1, z_1) \stackrel{\text{def}}{=} p_{out1}(x_1|z_1) \quad \text{and} \quad p_2(z_1, z_2) \stackrel{\text{def}}{=} p_M(z_1|z_2) \quad \text{and} \quad p_3(x_2, z_2) \stackrel{\text{def}}{=} p_{out2}(x_2|z_2),$$

with p_{out} a probability distribution that will be specified in each one of the parts explicitly later.

a) In this part let $p_{out1} = p_{out2} = p_{out}$ be a probability distribution on the set $\{1, 2, 3\}$ defined as

$$p_{out}(1|Hot) = 0.2, \quad p_{out}(2|Hot) = 0.4, \quad p_{out}(3|Hot) = 0.4,$$

$$p_{out}(1|Cold) = 0.5, \quad p_{out}(2|Cold) = 0.4, \quad p_{out}(3|Cold) = 0.1.$$

Compute the each one of the following entities, presented in (a1) and (a2), in two ways. For each one of the these,
First, go through an exact computation using the definition, and
second, use a suitable iterative algorithm you have learned in the class.
Then, eventually, compare your outcomes of these two methods in each case!
The entities to be computed are:

a1) The partition function

$$p(x_1 = 3, x_2 = 1) = \sum_{z_1, z_2} p(x_1 = 3, x_2 = 1, z_1, z_2).$$

a2) The most probable pair (z_1, z_2), assuming that one has observed an output $(x_1 = 3, x_2 = 1)$.

b) In this part let p_{out1} and p_{out2} be normal distributions with the means μ_i and variances σ_i for $i \in \{1, 2\}$, respectively. Assume that we have fully observed the following data in 6 independent trials,

$$(x_1 = 2.5, x_2 = 1.5, z_1 = Hot, z_2 = Cold), \; (x_1 = 1.2, x_2 = 3.1, z_1 = Cold, z_2 = Hot),$$

$$(x_1 = 2, x_2 = 2, z_1 = Hot, z_2 = Cold), \; (x_1 = 2.9, x_2 = 0.8, z_1 = Hot, z_2 = Cold),$$

$$(x_1 = 2.8, x_2 = 1.3, z_1 = Hot, z_2 = Cold), \; (x_1 = 2, x_2 = 2.7, z_1 = Cold, z_2 = Hot).$$

Estimate the best parameters (mean and variance in each case) for the output normal distributions, matching such observations using an ML algorithm.

Now, for Parts (c), (d) and (e) let,

$$p(x_1, x_2, z_1, z_2) = e^{-H(x_1, x_2, z_1, z_2)},$$

in which

$$-H(x_1, x_2, z_1, z_2) = \beta(x_1 z_1 + z_1 z_2 + x_2 z_2) + h z_1,$$

and

$$p_n(x_1, \ldots, x_{2n}, z_1, \ldots, z_{2n}) \overset{\text{def}}{=} \prod_{1 \le i \le n} p(x_{2i-1}, x_{2i}, z_{2i-1}, z_{2i}) e^{\beta z_{2i} z_{2i+1}}.$$

with

$$\forall\, 1 \le j \le 2n, \quad \{x_j, z_j\} \subseteq \{1, -1\}.$$

c) Approximate the log-partition function

$$\log Z_2 \overset{\text{def}}{=} \log \sum_{(x_1,\ldots,x_4,z_1,\ldots,z_4)} p_2(x_1, \ldots, x_4, z_1, \ldots, z_4),$$

for $\beta = 2$ and $h = 1$ (better approximations get higher credit!).

d) Write down the BP iterative dynamics for Z_2. Analyze this dynamic and determine its approximation factor by computing the exact values. Then, for the next part find the best *linear* approximation for this *nonlinear* dynamics (as a further approximation stage to simplify the algorithm) and explicitly find the corresponding matrix. Try to determine the approximation factor of this linear dynamics (better results get higher credit!).

e) Compute the first two terms of the loop series of

$$\log Z_1 \overset{\text{def}}{=} \log \sum_{(x_1,x_2,z_1,z_2)} p(x_1, x_2, z_1, z_2),$$

for $\beta = 2$ and $h = 1$. In this computation clearly write down your *normal* factor graph.
Then, try to linearize the sum of the first two terms of the loop series and compare your outcome with Part (d).

f) Compute the first four terms of the high temperature (i.e. β close to zero) expansion of

$$\log Z_n \overset{\text{def}}{=} \log \sum_{(x_1,\ldots,x_{2n},z_1,\ldots,z_{2n})} p_n(x_1, \ldots, x_{2n}, z_1, \ldots, z_{2n}),$$

when $h = 0$ and n is large.