CS 476/676, Spring 2016
Problem Set #3a: Posterior Inference
Version 1.0
Due by **11:59pm** on **Wednesday, March 30**

Contents

1 Instructions

1.1 What You are Given

We have provided text files contained in `hw3a-files.zip` which contain the input files you will use.

1.2 What to Hand In

All of your submission files should be handed in as a single archive named `hw3a-username.zip`, where `username` is replaced with your JHED ID. This zip archive will include:

- The code you write, your program as an executable, and various input files that you create. Instructions for what to hand in are given under the section headings titled **Deliverables**.
- Written answers to questions should go in a single PDF document named `writeup.pdf`. Questions that require a written response are given under the section headings **Analytical Questions** or **Empirical Questions**. We recommend using LaTeX to typeset your writeup.

1.3 Submission Policies

Please note the following:

- **Collaboration:** Please work in groups of size 2 people. The homeworks are a way for you to work through the material you're learning in this class on your own. But, by working in a group, and debugging each other's solutions, you'll have a chance to learn the material in more depth. The recommended format for tackling these problem sets is the following. Write a high level sketch of the solution for all of the problems on your own. Meet as a group to brainstorm your solutions and converge on a solution as a group. It is important that you have a good understanding of how you'd have approached the problem independently before discussing your solution with the other group members. Developing this intuition will serve you well in the final exam where you will be required to work on your own. Pursuant to your group meeting, write up the solutions on your own. Thereafter, meet as a group to clean up and submit a final write up as a group. By now, each of you should have a solid understand of the concepts involved, and by meeting as a group, you've had a chance to see common ways in which one can make mistakes. Submit your final solution as a final writeup for the group. Your submission should include the names of every team member. Also, name your file as `hw3a-username1-username2.zip`.
- **Late Submissions:** We allow each student to use up to 3 late days over the semester. You have late days, not late hours. This means that if your submission is late by any amount of time past the deadline, then this will use up a late day. If it is late by any amount beyond 24 hours past the deadline, then this will use a second late. **If you jointly submit an assignment as a team, then every team member will lose late days if the assignment is submitted late.** If you collaborate with team members but independently submit your own version, then late days will only apply to you.

1.4 How to Hand In

As with the previous homework, one person per group should submit the assignment on Piazza. A private note to the instructors should be submitted with the subject *Submission 3a from* `<list`

`of names of team members>` with the .zip submission as an attachment. The note should be submitted to the `submission3a` folder.

2 Variational Inference on a Simple Network [40 points]

In this question, we will explore the difference between mean-field and structured mean-field variational inference. We want to perform inference in the network shown below, where each node has values in $\{1, 2\}$ with the following conditional probability distributions (the conditional distributions have already been implemented for you in the file `probs.R`). We are interested in computing the marginal distribution over E and F.

For this problem, we have provided starter code in three files: `probs.R`, `inference.R`, and `demo.R`. `probs.R` contains the conditional probability distributions (implemented as functions) along with some utility functions used in `inference.R`. `inference.R` contains two functions: `mean.field.inference` and `struct.mean.field.inference`. Both of these functions are only partially implemented. You will complete the implementation using update equations that you derive in the following questions. Finally, `demo.R` is a script that can be run by typing `Rscript demo.R` at the command line. Initially, this will only print the true marginal distribution over E and F and will fail with an error message indicating that update functions have not yet been implemented. Run this script once to see what the true marginal looks like.

$P(A = 2)$	0.5
$P(B = 2\|A = 1)$	0.1
$P(B = 2\|A = 2)$	0.9
$P(C = 2\|A = 1, B = 1)$	0.1
$P(C = 2\|A = 1, B = 2)$	0.5
$P(C = 2\|A = 2, B = 1)$	0.5
$P(C = 2\|A = 2, B = 2)$	0.9
$P(D = 2\|B = 1)$	0.4
$P(D = 2\|B = 2)$	0.6
$P(E = 2\|C = 1, D = 1)$	0.1
$P(E = 2\|C = 1, D = 2)$	0.9
$P(E = 2\|C = 2, D = 1)$	0.3
$P(E = 2\|C = 2, D = 2)$	0.7
$P(F = 2\|D = 1)$	0.1
$P(F = 2\|D = 2)$	0.9

Table 1: Conditional probability distributions for the simple network.

2.1 Empirical Questions [40 points]

a. **[10 points]** Derive the mean-field variational update equations for this network. Start with the energy functional, and explicitly derive the update equations for each of the marginal variational distributions (e.g. $Q(A)$).

b. **[10 points]** Implement the update equations in the function `mean.field.inference` in `inference.R`. Note that you only need to implement the $\{$`a,b,c,d,e,f`$\}$`.up` functions; everything else has already been written for you. Once you have completed the implementation, re-run `demo.R`. The script should print the marginal over E and F obtained using mean-field inference, and will display the KL divergence between the true and approximate marginal. Answer the following questions:

- Does this look like a reasonable approximation?
- What does the KL divergence mean in this case?

c. **[10 points]** To improve the approximation, we can use structured mean-field variational inference. Instead of using a completely factored joint over A, B, C, D, E and F, we will use two factors (A, B, C) and (D, E, F) (i.e. we have two variational distributions $Q(A, B, C)$ and $Q(D, E, F)$). Derive the update equations for this new factorization. Again, begin with the energy functional and explicitly derive the update equations for each factor.

d. **[10 points]** Implement the update equations in the function `struct.mean.field.inference` in `inference.R`. You only need to complete the functions `abc.up` and `def.up`. Re-run `demo.R` after completing the update functions. Again, the script should print the mean-field and structured mean-field approximate marginals below the true marginal along with the KL divergence for each. Answer the following questions:

- Does this look like a better approximation than mean-field?
- What does the KL divergence mean in this case?
- Give a brief explanation (2-3 sentences) for why structured mean-field performs better for this Bayesian network.

The functions you need to implement only require assigning variables and calling basic math functions; we expect you to be able to do this even if you have no previous experience with R. If you need an introduction to R, we recommend looking at `http://cran.r-project.org/doc/manuals/R-intro.html`.

Deliverables Submit the script `inference.R` with your implementation of the required functions.

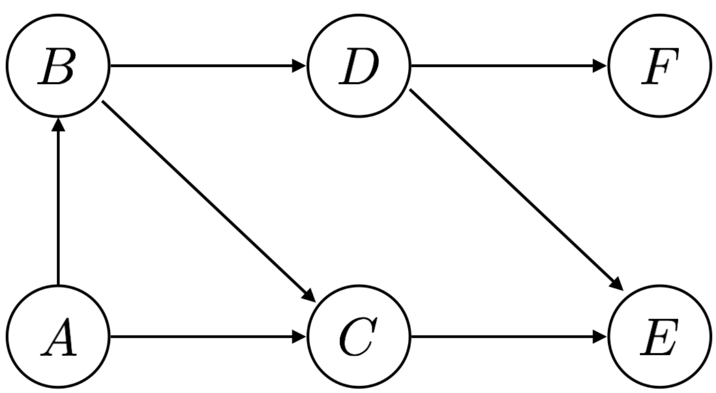

Figure 1: Simple network