CSC 535 – Probabilistic Graphical Models

Assignment Seven

Due: 11:59pm (*) Friday, Nov 9.

(*) There is grace until 8am the next morning, as the instructor will not grade assignment before then. However, once the instructor starts grading assignments, no more assignments will be accepted.

Weight is about 7 points

This assignment and the next can be done in groups of 2-3. You should create one PDF for the entire group, but everyone should hand in a copy of the PDF to help me keep track. Make it clear on the first page of your PDF who your group is. Group reports are expected to be higher quality than individual ones.

The purpose of this assignment is to learn about exact inference by implementing the sum-product and max-sum algorithms. We will also consider one way to visualize how well we can infer models from data.

Deliverables

Deliverables are specified below in more detail. For the high level perspective, you are to provide a program to output a few numbers and to create figures. You also need to create a PDF document that tells the story of the assignment including output, plots, and images that are displayed when the program runs. Even if the question does not explicitly remind you to put the resulting image into the PDF, if it is flagged with **($)**, you should do so. The instructor should not need to run the program to verify that you attempted the question. See

http://kobus.ca/teaching/grad-assignment-instructions.pdf

for more details about preparing write-ups. While it takes work, it is well worth getting better (and more efficient) at this. A substantive part of each assignment grade is reserved for exposition.

Sex in flatland (the sequel)

This assignment is based on the previous assignment. Please consult that assignment for the specification of the MQL model.

1. (++++) Implement a sum-product algorithm to determine the sex of the MQL based on the observed data. Your algorithm should provide the posterior probabilities for male versus female **($)**. Now suppose that you have to declare whether an MQL is male versus female. The natural thing to do is to provide the maximally probable one, as we do not have a specific risk function in mind. Test the accuracy of doing so on enough generated data so that the accuracy estimate is stable **($)**. Finally, to get used to putting such numbers in perspective, you should tell the reader what "chance" performance is. (Hint. It is not 50%) **($)**.

   Consulting the oracle. *Once you have this running to your satisfaction, you may email the instructor with your estimate of accuracy. The instructor will let you know if you are in the range of expected values.*

   Tips. *Your code need not be particularly long, but some students in previous years have reported that they found this question difficult. You should budget some quality time for this, and start early so that you can get help, and/or work on it over multiple sessions.*

   I recommend making a factor graph, and keeping it handy while you code. Also, getting this right depends on a good answer to some of the previous assignment so be critical of your generative model. Debugging in stages is highly recommended in general, and is more critical when the algorithm is non-trivial or, at least, new to you. You should challenge yourself to figure out ways to test small portions of the code.

   You may wonder how you can validate the code as a whole. Two suggestions: 1)You will need to condition on the observed nodes to find the posterior. However, if you do not condition on anything, and marginalize out everything, then the answer should be 1. Check this! 2) I have specified distributions that separate male and female to some extent, but appropriately changing them will make males and females more or less distinguishable. You should verify that making them progressively more distinguishable, increases performance.

   It is possible to provide observations that lead to zero probability. This is not the case if you use data generated from the model. But if you create observations for code development and debugging you want to be sure that they are feasible.

2. Generate a plot indicating the percentage of true males as a function of probability of males provided by your program, in suitable increments (e.g., 0-10% chance of males, 10-20%, etc.), and comment on the results **($)**. This task is easier to understand if you realize that the goal is to visualize the extent to which your program being more sure about male (higher estimated probability) correlates with the true sex is male.

   More explanation. *For (sufficient) generate data, study the examples where the sum-product algorithm says the probability of male is in each of collection of ranges (e.g., 0-10%, i.e., 5±5%, which you might plot as 5% on the x-axis, 10-20%, i.e., 15±5%, etc.). Now compute the empirical probability of male from the ground truth. If you algorithm spoke the truth, maybe about 5% would be male. Regardless, this empirical estimate would form the y-axis value for the this range. As the range center increased, one might hope that the true percentage also increased, which is what you are investigating.*

3. (++) Implement a max-sum algorithm to determine the maximal configuration of the MQL hidden variables. Compute this for at least 10 generated measurement sets, and plot the observed version of the MQL, the inferred configuration of the MQL, and the true MQL. Hopefully it natural to have rows of three images, ordered observed, inferred, and true.

4. Similar to what you did in question (1) for sum-product, test the accuracy of using the max-sum algorithm on enough generated data so that the accuracy estimate is stable **($)**. Compare the accuracy of max-sum and sum-product, and comment on this **($)**.

What to Hand In

Hand in the PDF file hw7.pdf with the story of your efforts into D2L. If you wrote code for this assignment (you should!), hand in a program hw7.<suffix> (e.g., hw7.m if you are working in Matlab). If you are working in groups, the code and the PDF can be the same for each person in the group, but each member should hand in a copy to help me keep track of things. PDFs from groups are expected to be of higher quality than individual ones.