

STAT 535 Homework 7
Out November 18, 2011
Due November 29, 2011

©Marina Meilă
mmp@stat.washington.edu

Problem 1 - Max Propagation implementation

Submit code for all question of this problem through the Assignments web site. You are required to implement MAXABSORB as a general purpose algorithm taking tables as parameters, in a representation of your choice. It is OK to special case the rest of the assignment by hand-coding calls to MAXABSORB. Please comment or write “readable” code.

In addition to submitting the program, the problem calls for answers to a few questions and printing of results. These should be returned on paper or in a separate .pdf file.

a. Implement the function `MAXABSORB(ϕ_{source} , ϕ_S , $\phi_{\text{destination}}$, I_{source} , $I_{\text{destination}}$, isCollect)`.

In the above, $\phi_{\text{source, destination}}$ will represent as usual the clique potential tables. The tables $I_{\text{source, destination}}$ represent the indicator variables for the argmax, that is, for the local variable configuration that achieves the maximum. The variable `isCollect` is a Boolean variable, which indicates the direction of the propagation. This is used for the operations with the I tables. If `isCollect=1`, the I_{source} is set. If `isCollect=0`, then we are in the backtracking stage: $I_{\text{destination}}$ absorbs from I_{source} .

Note: make sure that you only set a single value of I_{source} to 1 for each configuration in the separator. (Only code required here.)

b. Consider the same junction tree as in Homework 6, given by the table in `hw6-ptables.dat`. Read the tables and create a calibrated junction tree structure representing this distribution. Do all this in code, and do not print anything for this question. (This is the same as **1.b** from Homework 6.)

c. Run the MAXPROPAGATION algorithm on the above junction tree and obtain $\max_{\Omega_V} P_V$ and $\text{argmax}_{\Omega_V} P_V$.

More precisely: Print $\max_{\Omega_C} P_C(x_C)$ for each clique in the junction tree. These values must be equal, as they all represent $\max_{\Omega_V} P_V$. Print the I tables after

COLLECT, then again at the end of max propagation, and extract from each of them the configuration indicated by the unique 1 in the table. Derive which is the global configuration $x_V^* = \operatorname{argmax}_{\Omega_V} P_V$.

Problem 2. - IPF with tied parameters



Consider the MRF $D-C$ where all variables are binary and $\phi_{AB} = \phi_{CD} = \phi$ and $\phi_{BC} = \phi_{AD} = \psi$. The difference between this model and the example in Lecture 12 is that the parameters for different potentials are constrained to be equal. We will derive the IPF update for this model.

- a. Write the expression of the log-likelihood $l = \ln P(\text{data})/N$ for this model. What are the sufficient statistics? Can they be interpreted as a probability distribution?
- b. Calculate the gradient of the log-likelihood obtained in a and show that it depends on difference(s) between data marginals and model marginals.

[c. – **Extra credit**] Implement IPF for this model. Note that for this you need to implement a form of inference in the MRF (for example by constructing a junction tree and calculating the model marginals from it).

Problem 3 – IPF in a decomposable model

Prove that in a junction tree, i.e. in a MRF over a chordal graph, there is a sequence of clique updates so that IPF converges in a finite number of steps.

More precisely, use the rooted tree factorization of the junction tree, which has no explicit separator potentials. (The statement is true for other j.t. parametrizations as well, but one would have to appropriately re-define the IPF algorithm.)

Assume that the initialization is compatible with some $P_V^0 > 0$ joint distribution over the variables in V . You must show that the algorithm finds the correct values after a finite number of steps from any such initialization. A “step” here means updating all the parameters of a single clique, i.e. updating ϕ_C for a single C . You can prescribe in the algorithm the order in which the updates should be made.

[**Optional, extra credit:** What would happen, if the initialization was not compatible with a P_V^0 , but an initialization with arbitrary potential values $\phi_C(x_C) > 0$? Can you modify the IPF algorithm, so that it converges to the correct j.t. parametrization for this initialization?]

[Problem 4 – Gaussian Junction Trees EXTRA CREDIT]

- a.** Write the ABSORB function and the Junction Tree algorithm for a gaussian graphical model. Assume that the clique and separator potentials are given by their parameters μ_C, Σ_C . What does it mean for a Gaussian j.t to be calibrated?

- b.** Write the MAXABSORB function and the Max Propagation algorithm for a Gaussian junction tree.