1) I am hoping that the variational expectation of the Cho et al paper for the beta parameter is wrong:

2) I cannot see much sparsity in the network I generate, it is actually quite well connected, but I am following the generative process. I was wondering how can we produce a sparser network using the same process. This is usually achieved when the hyperparameter α_k's for the membership vector are set to small values.

3) I have to recheck the derivation of the lower bound. The increase in log likelihood is guaranteed. But I see from time to time very small glitches. I have two points of doubts, one in the first line about link probability, and the other one about the η, τ, in the derivations of the parameters especially.

4) I need to reconfirm the way I implement my sums regarding the derivations for the variational parameter updates. Especially the ones I go through links, and the ones I go through neighbors of each individual. The first has no double counts, but the second one has. The derivations don't imply any other ways. But I need to make sure about that.

5) If we are doing Variational EM, then this is all just the E-step, and we need to estimate the model parameters(hyperparameters) as well, so we need a way to estimate η_0, η_1 and also α. This requires the same maximization of the ELBO with respect to the model parameters, fixing the variational parameters. But the expressions are not analytical, I will need help with that. So in other places there are some suggestions for example for using Newton-Raphson method. Besides after updating the variational parameters, we should be able to reconstruct the actual parameters of the model in the mean field factorized distribution.I use some intuitive estimate for the parameters, such as θ and β. I also need to check this. For example we have to be able to reconstruct the Z from ϕ, or θ from γ, and β from τ. For $\theta_{a,k} = \dfrac{\gamma_{a,k}}{\sum_l \gamma_{a,l}}$, and for $\beta_{k,k} = \dfrac{\tau_{0,k}}{\tau_{0,k} + \tau_{1,k}}$. But I am not clear whether these are reasonable, plus I don't see much good output.

For now I am fixing the model parameters skip the M-step, and only doing the inference on the variational parameters, assuming I have the correct model parameters.

6) the problem of the smart/fast initializations of the global parameters still remains, for now I skipped it, because I expect the random initializations also should work. But it would be necessary to do this as it offers a good close enough number for K. But for now I use the correct number K of the ground truth.

7) Should I remove the parts in computing the ELBO, that are independent of the changes in the parameters, for example if I am updating the variational parameters where model parameters are untouched, can I skip those segments in the ELBO that are entirely a function of model parameters?

8) The ELBO behaves alright, but the values I am not sure if makes sense or not.

9)I do not exactly know how should I check the results. What should be the measure for evaluation? Should I use the estimated parameters to create links and memberships based on them to see how well I am doing?

VISUALIZE HELP

10) specify initializations

11)The data given, is directed however, this will change the updates