Advanced Probabilistic Machine Learning and Applications: Exam (29-09-2021)

1 Exercise 1: Variational inference and topic modeling

1. What is the main idea of Variational Inference (VI)?
2. How is this different than Gibbs sampling?
3. Describe the Latent Dirichlet Allocation (LDA) for topic modeling. In particular, i) describe what are the variables involved and what is the main idea of the model; ii) write the generative process drawing the various variables: β_k (word proportion), θ_d (topic proportion), z_{dn} (topic assignment) and w_{dn} (word).
4. Describe Mean Field Variational approximation for the LDA and describe what would be the variational distributions for the latent variables.
5. Consider the variational distributions proposed in the previous question. Are these the optimal ones? Discuss about this.
6. What are the main assumptions and possible limits of LDA?
7. Describe the CAVI (coordinate ascent VI) procedure.

2 Exercise 2: matrix factorization

1. Describe Poisson matrix factorization and its main assumptions. It can help to start by defining what the data A, a matrix with entries A_{ij}, can represent in example applications.
2. State various inference approaches and estimation methods that can be used to infer the parameters. Mention their differences and similarities (if any).
3. One main parameter is the affinity matrix with entries c_{kq}. Derive the Maximum Likelihood estimate $\hat{c}_{kq}$ of this. Hint: write the Poisson likelihood[1] and use Jensen's inequality[2] with a variational distribution q_{ijkq}.
4. Describe possible constraints on the parameters and connections with regularization.
5. Consider now matrix factorization (not the Poisson one). How does this differ from the Poisson case described above?
6. Describe possible approaches to tackle matrix factorization and their connections (if any).
7. (extra) What could be possible problems of the basic approaches? How would you address it?.

[1] A Poisson distribution for a variable x and parameter λ is $P(x|\lambda) = e^{-\lambda}\lambda^x/x!$.

[2] $\log\left(\sum_k x_k\right) \geq \sum_k q_k \log \frac{x_k}{q_k}$, where q_k are any probabilities satisfying $\sum_k q_k = 1$ and x_k are any set of positive numbers.