CS599-Advanced Topics in Neural Computation and Statistical Learning

Questions for Jordan&Bishop Chapter 11

1. Why EM?
2. What does it mean to simplify statistical inference by introducing latent variables?
3. Why is the likelihood function a marginal probability when the latent variables are not observed?
4. What is meant by “decoupling the estimation problem?”?
5. What is the incomplete and the complete data likelihood, and how are they related?
6. How is the expected complete log likelihood defined?
7. How is EM generally derived?
8. What is Jensens’ Inequality, how is it derived
9. Why is $q(z|x)$ treated as independent of the parameters □?
10. What is the proof of EM convergence?
11. What is the KL-divergence?
12. What is the KL-view of EM?