CS599-Advanced Topics in Neural Computation and Statistical Learning

Questions for Jordan&Bishop Chapter 10

1. What is a latent variable model?
2. What is the use of introducing a latent variable?
3. What is a "devide-and-conquer" approach to statistical modeling?
4. What is the difference between clustering and classification?
5. What is the generic formulation of an unconditional mixture model, and what are the distributions of the random variables?
6. Are mixture model solved traditionally by Bayesian statistics?
7. Why is a "mixture of Gaussian" model harder to estimate than a single Gaussian, and what methods can be used to do this?
8. Why is the EM algorithm of central importance in statistical learning?
9. For which kind of learning framework was EM created?
10. What is the k-means algorithm, and how does it work?
11. What is the "chicken-and-egg" problem in k-means?
12. What is the optimization view of k-means, i.e., which optimization criterion is used?
13. How does EM work for mixture models, in comparison with k-means?
14. Why is EM a soft-assignment algorithm, and k-means a hard-assignment one?
15. What is the basic structure of EM, i.e., what is the E-step and the M-step in the context of mixture models?
16. How do you initialize EM for mixture models, and how sensitive is EM to initialization?
17. Demonstrate the EM update corresponds to stationary points of the log likelihood function.
18. What is the expected complete log likelihood, how is it formulated for unconditional mixture models?
19. What is the "expected complete data log likelihood?
20. How is EM derived from the complete data log likelihood view?
21. What is the true definition of the E-step in EM?
22. What is the true definition of the M-step in EM?
23. Derive EM update from the complete data likelihood point.
24. What are the new elements added in conditional mixtures?
25. What is the conditional mixture model formulation?
26. Formulate a conditional mixture with logistic regression and Gaussian models as classifiers, and discuss this in terms of a "gating network".