

COMPSCI 271 - Machine Learning

Homework Due 11/20/07

1 Clustering and EM I

Consider the EM algorithm, as presented in the textbook, but modified so that the **E** step is a “hard” E step. By “hard” we mean that instead of assigning responsibilities between 0 and 1, the E step assigns $\gamma(z_{nk}) = 1$ for the value of k that maximizes the equation 9.23, and $\gamma(z_{nk}) = 0$ for all other k . Show that this version of EM must converge in a finite number of iterations.

2 Clustering and EM II

Suppose you are using EM to find a Gaussian mixture model and that the initial cluster centers and covariances are chosen to be identical. Describe precisely what happens in the first E step and the first M step. Do you think this is a good starting point?

3 Clustering and EM III

Suppose your professor uses the following method for assigning grades: He uses a Gaussian mixture model to learn three clusters of grades, then assigns the cluster with the highest mean an A, second highest mean a B, and third highest mean a C. Give at least two reasons why this might not be a fair grading system.

4 Variational Methods

Do problem 10.30.

5 Sampling I

Do problem 11.2.

6 Sampling II

Do problem 11.4. (Show that the Box Muller method generates samples from the Gaussian distribution.)

7 Sampling III

Do problem 11.10. (This is mostly to give you more practice doing proofs.)

8 Sampling IV

Do problem 11.13.

9 PCA

In class, we made the comment that minimizing the reconstruction error for PCA was equivalent to maximizing the variance of the reconstructed data. 1) Write down an expression for the variance of the reconstructed data. 2) Show that maximizing your expression from (1) is equivalent to the minimization demonstrated in class.