Homework 6, EE 527, Spring ~~2008~~

~~Due March 27 (for both EDE and on-campus students)~~

1. Derive the EM algorithm for a Gaussian mixture model. Consider scalar data $x_1, x_2, \ldots x_N$ which is i.i.d. and distributed according to a $p = 3$ class Gaussian mixture with mixture probabilities $\alpha_k, k = 1, \ldots p$, means $\mu_k, k = 1, \ldots p$, and variances $\sigma_k^2, k = 1, \ldots p$, **all unknown**, i.e.

$$x_i \sim \sum_{k=1}^{p} \alpha_k \mathcal{N}(\mu_k, \sigma_k^2), \quad \forall i = 1, 2, \ldots N$$

   You can use the paper posted on the class webpage (A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models, by Jeff Bilmes), but understand the method carefully and rewrite it. Also fill in the gaps in the derivation.

   Practice ungraded parts (I will suggest doing at least one)

   (a) Implement the EM algorithm that you derive. Also, try to implement Newton Raphson for the same problem and see the difference in results. To obtain any reasonable conclusions you need to do a Monte Carlo simulation test.

   (b) Implement a alternating minimization algorithm for computing the mixture parameters. This is the exact same problem as K-means clustering.

2. Consider i.i.d. data $x_1, x_2, \ldots x_N$ where

$$x_n = \sum_{k=1}^{p} r_k \cos(\omega_k n + \phi_k) + w_n, \quad w_n \sim \mathcal{N}(0, 1) \tag{1}$$

   You need to compute the ML estimates of $\omega_k, r_k, \phi_k, \ k = 1, 2, \ldots p$. Rigorously show that you can convert the above problem into a problem where only p parameters need to be estimated using Newton Raphson (note: I have explained the idea in class for $p = 1$, but now use the theorems in the book and anything else you need to rigorously show that this holds).

   (a) Implement Newton-Raphson to compute the ML estimates.

   (b) **Practice Ungraded:** Also, directly implement Newton-Raphson for all the parameters, with the same number of iterations that the first part used. See what you get. Do a full Monte Carlo simulation to show that the first part gives "better" results. Define "better" as you feel like.

3. Refer the to Recursive Least Squares algorithm proof given in the Least Squares Estimation handout. I have left some gaps there. Fill in the gaps, understand the whole thing and rewrite.

4. Kay's book: Problems 8.24, 8.26, 8.27 (skip the Newton Raphson part)

5. **Practice problems (I will suggest doing at least two):** 8.4, 8.12, 8.28, 8.29