

*Instructions:* Please put all answers in a single PDF with your name and NetID and upload to SAKAI before class on the due date (there is a LaTeX template on the course web site for you to use). Definitely consider working in a group; please include the names of the people in your group and write up your solutions separately. If you look at any references (even wikipedia), cite them. If you happen to track the number of hours you spent on the homework, it would be great if you could put that at the top of your homework to give us an indication of how difficult it was.

### Problem 1

#### *MCMC for a Gaussian Mixture Model*

Let

$$x, z, \mu, \Sigma, \pi \sim p(x|z, \mu, \Sigma)p(z|\pi)p(\pi) \prod_{k=1}^K p(\mu_k)p(\Sigma_k)$$

for a Gaussian mixture model with  $K$  mixture components, as in 24.2.3 of the Murphy book. We will keep the model identical to the model in the book, with a slight modification: let the cluster-specific mean parameters  $\mu_k \sim \text{Unif}(0, 5)$  have a uniform distribution (over 0, 5) rather than a Gaussian distribution. Download the `HW6_mixture.txt` data from SAKAI for this problem. This is a univariate Gaussian observation, so the inverse Wishart distribution can be replaced by the simpler inverse Gamma (conjugate) distribution for the component-specific variance terms. Here, you can set  $K = 2$ .

- (a) Write out a possible Metropolis-Hastings step for the  $\mu_k$  parameters, replacing the Gibbs sample for these variables in 24.2.3. What is the proposal distribution (hint: make it a simple 'step' given the current value of the mean parameter)? What is the MH acceptance probability?
- (b) Why do we choose to perform MH here instead of a Gibbs sample step?
- (c) Implement MCMC for GMMs in R. Show your code. How many iterations of Burn-In did you run? How many iterations of sampling did you run? How did you initialize your parameters?
- (d) Show the log likelihood trace for three different runs of the sampler starting at three different points on the data you downloaded.
- (e) Plot a histogram of the posterior samples for each mean parameter for a single run (after burn-in). Write one sentence about what this means. Did label switching occur?
- (f) How might you choose a single estimate for the component-specific means and variances? What are those values on the data?