

Machine Learning Module

Week 6

Laboratory Exercise, Week 6

Probability Density Estimation

Mark Girolami

`girolami@dcs.gla.ac.uk`

Department of Computing Science

University of Glasgow

February 5, 2006

# 1 Probability Density Estimation

Download the Matlab file `Lab_6_EM_Data.mat`. You will find a data matrix  $\mathbf{X}$  which should be used to run the EM algorithm to estimate the parameters of a Gaussian Mixture model. No details are available as to the actual density of the data so it is your job to assess the number of mixture components which will be most appropriate for this data. Use the data stored in the matrix  $\mathbf{Xt}$  to compute the *test data* likelihood under each mixture model with varying numbers of components ranging from say 2 to 10. What is the optimal number of mixture components?