CS-ECE4720/7720 Spring 2015
Introduction to Machine Learning and Pattern Recognition

Homework Assignment 3

due Thursday 3/19/2015

Problem 1. Textbook p. 140, Prob. 1 (a) & (b).

Problem 2. Textbook p. 148, Prob. 25

Problem 3. Two datasets have been generated from 2-D Gaussian sources $N(\underline{\mu}, \Sigma)$. Implement the maximum likelihood estimation equations for $\underline{\mu}$ and Σ (Eqs.(18) &(19) on page 89) in MATLAB and report your results as specified below in Part I and Part II.

Part I. Report your results on the dataset GDdataMLE1.txt which is with $\underline{\mu} = [1 \quad -1]^T$, and $\Sigma = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

a) Estimate $\underline{\mu}$ and Σ from the first 10 data samples;
b) Estimate $\underline{\mu}$ and Σ from the first 100 data samples;
c) Estimate $\underline{\mu}$ and Σ from the first 1000 data samples;
d) Estimate $\underline{\mu}$ and Σ from the first 10000 data samples;
e) Compute the parameter estimation errors for the above four cases by using the following two error measures:

   Measure 1: $\varepsilon = \left\| \hat{\theta}_{MLE} - \theta_{true} \right\|$, where $\theta = [\mu_1, \mu_2, \sigma_1^2, \sigma_{12}, \sigma_2^2]^T$.

   Measure 2: $\varepsilon_\mu = \frac{1}{\sqrt{2}} \left\| \hat{\theta}_{\mu,MLE} - \theta_{\mu,true} \right\|$, $\varepsilon_\Sigma = \frac{1}{\sqrt{3}} \left\| \hat{\theta}_{\Sigma,MLE} - \theta_{\Sigma,true} \right\|$, with θ_μ and θ_Σ consisting of the first two and the last three components in θ, respectively.

   Put the error results in a table and discuss your observations.
f) For the case c), plot in one figure the first 100 data samples and the 2D contours of the estimated Gaussian pdf at the level of $\frac{1}{2\pi|\Sigma|^{1/2}} e^{-1}$ (see the PlotContourExamples for examples of plotting the contours of a density function).

Part II. Following the five steps of Part I, report your results on the dataset GDdataMLE2.txt which is generated with $\underline{\mu} = [1 \quad -1]^T$, and $\Sigma = \begin{bmatrix} 2 & -1.9 \\ -1.9 & 2 \end{bmatrix}$.

Problem 4 Implement MLE and MAP estimations for the mean parameter of a Gaussian density $N(\mu \ \sigma^2)$, where $\sigma^2 = 2$, $\mu_0 = 2.2$, $\sigma_0^2 = 0.25$.

Use the dataset GDdataMLEMAP.txt to evaluate the estimation accuracy for the two methods in the following way: perform 20 estimation experiments by using the first 2(2k-1) samples, for k = 1, 2, …, 20.

a) Report the MLE and MAP estimated mean parameters for the twenty cases.
b) Given the true mean parameter = 2, compute the MLE and MAP estimation errors for each of the twenty cases, and plot the error curves in one figure (use two different colors for MLE and MAP). Discuss your observations on the error curves.
c) Plot the posterior probability density function of $\mu \sim N\left(\mu_N, \sigma_N^2\right)$ for the cases of k=1, 10, and 20. Discuss your observations.

Problem 5

The Iris dataset was created by R. A. Fisher and has been well studied in machine learning: iris.names.txt and iris.data.txt.

There are three classes in the Iris dataset: Iris Setosa, Iris Versicolour, and Iris Virginica. For each class, take the first 10 samples and the last 10 samples as the test data and the middle 30 samples as the training data (each class has 50 samples).

Experiment 1

For this experiment, we'll use one multivariate Gaussian density with a diagonal covariance matrix to model the 4-dimensional data distribution of each class.

a) Implement the maximum likelihood estimation of the mean vector and the covariance matrix (you can implement a general function in Prob. 3 and use it here);
b) Compute the MLE estimates of the mean and the covariance matrix for each class and report the 3 mean vectors and 3 covariance matrices.
c) Use maximum likelihood classification to classify the 60 test samples, and tally your classification errors in a confusion table as shown below :

Classified class / True class	Setosa	Versicolour	Virginica
Setosa			
Versicolour			
Virginica			

Experiment 2

For this experiment, we'll use one multivariate Gaussian density with a full covariance matrix to model the 4-dimensional data distribution of each class, and do a) b) and c) as specified in Experiment 1.

Attach your program code.