UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
Department of Electrical and Computer Engineering

ECE 544NA Pattern Recognition

Homework 6

Fall 2013

Assigned: Thursday, October 17, 2013 Due: Tuesday, October 29, 2013

Reading: NNPR Chapters 5, 6, 7

Problem 6.1

The Bayes rule for some loss $L(y, f(x))$ is said to be $f(x)$ that minimizes the expected loss:

$$f_B(x) = \arg\min_{f(x)} E_{X,Y}[L(y, f(x)]$$

We saw in class that when $L(y, f(x)) = |y - f(x)|^2$, $f_B(x) = E[Y|X]$, the conditional mean. Find the optimal (Bayes) rule for $L(y, f(x)) = |y - f(x)|$.

Problem 6.2

In Homework 2, you implemented the expectation maximization (EM) algorithm for learning the parameters of a Gaussian mixture model (GMM). The k-means algorithm also appears to have an iterative procedure that resembles EM. Show that when the covariance matrices of the Gaussians are assumed to be ϵI, where I is the identity matrix, the EM algorithm for a GMM reduces to k-means clustering as $\epsilon \to 0$. Note: k is the number of clusters.

Matlab Exercises

Problem 6.3

Again, for the diabetes dataset, create a training set and a test set as you did in homeworks 4 and 5. Consider the single-layer neural network from HW 4 with the $tanh()$ activation function. As always, you are expected to write your own code, but you may reuse code from previous assignments.

(a) Incorporate Newton's method into your single-layer network from HW 4 and plot the error on the **training set** as a function of the iteration number.

(b) Use the Levenberg-Marquardt approximation for the (inverse) Hessian and again, plot the error on the **training set** as a function of the iteration number.

(c) Plot the error on the **training set** as a function of the iteration number for the simple gradient descent algorithm used in HW 4. Compare the convergence of Newton's methods to gradient descent. Please run all three on the same training dataset with the same initialization. Show the three graphs on the same plot (using *hold on* in matlab).

(d) Test your three trained models on the **test set** and comment on your results.