**ECE 595 Homework 9 Due: Dec. 5, 4 PM**

The unlabeled 2-D dataset from the file *pcadata1.mat* has one direction of large variation and one of smaller variation. (a) Perform principal component analysis analytically on this data, i.e., obtain the mean-removed data, covariance matrix, eigenvalues and eigenvectors. You may use the *eigen* or the *svd* function in Octave/MATLAB. Obtain the first principal component – projection along the eigenvector corresponding to the largest eigenvalue – and verify how much of the data variance is available in this component using

. Verify that this ratio is the same as , where , i = 1, 2, … *m* are the ordered eigenvalues (obtained from the *svd* function).

(b) Use a 2-2-1 layer neural network with a learning rate of 0.005 or smaller and iteratively obtain the first principal component. You may end up with the first principal eigenvector that is negative of the vector obtained in (a).

**Optional Exercise for 30 points**: Use the *svd* function on the overlapping half-moon data of Homework 7 (*hm7.mat*) and obtain the first principal component in 1-D. Note that you are now using labeled data; therefore, the labels must be carried over to the principal component for classification. (a) Run the perceptron LMS algorithm and determine if the data can be linearly separated. (b) Do polynomial features from the principal component work better in linear separation of the two classes? Determine using logistic regression first with  and next with and may also include , where *x1* is the principal component. Note that you may need to normalize if the polynomial data are of different range.