Homework 6
Posted Tue, Feb 26, 2008
Due Tue, March 4, 2008

Problem 1. (50 points)

1.1
Calculate the principal eigenvectors of the iris data set (all four), using the MATLAB routine princomp. Use the first four columns of the "iris-train.txt" data, preprocessed to have zero mean in each column. (Zero mean of the samples in all four dimensions.)
Record the eigenvectors and eigenvalues.
Be sure to check that P*P' = I, where P is the matrix of the eigenvectors and I is the identity matrix. P' is the transposed of P.

1.2
Write a MATLAB function to implement Sanger's Generalized Hebbian Algorithm (GHA), and calculate the principal eigenvectors of the "iris-train.txt" data (first four columns as above) using the GHA. Again, check how close P_{GHA}*P_{GHA}' is to I, or check periodically.

Compare the results obtained in 1.1 .and 1.2. Report your initial weight matrix, the number of iterations that were needed to converge, the learning rate, and the training history. The training history can be, for example, the matrix P_{GHA}*P_{GHA}' (how close it is to I), and you may also print the difference between old and new weights. (You will see that very small differences in the latter do not necessarily indicate converged state, as seen from the P_{GHA}*P_{GHA}'matrix.)

Write the implementation of the GHA similarly as you wrote your BP implementation in HW-5, i.e., write the GHA itself as a function, and call it from a main program where the I/O, pre- and post-processing, reporting are done.

(Suggestion: You may first set the number of output PE's to 1, to compute just the first principal eigenvector, in order to save time testing your code and to get a feel for the parameters.)