

Machine Learning Module

Week 7

Laboratory Exercise, Week 7

Principal Component Analysis

Mark Girolami

`girolami@dcs.gla.ac.uk`

Department of Computing Science

University of Glasgow

February 23, 2006

1 Principal Component Analysis

PCA has been shown to be a useful method of linear feature extraction and dimensionality reduction for 64×64 pixel images of faces. In Week 5 a collection of images of handwritten digits (3 & 8) were used in the Laboratory. Using these images perform PCA on the collection of 3's and then 8's. What features of the original digits do the images associated with each of the principal directions capture? Examine the eigenvalues associated with each principal direction how many components are required to obtain 10% reconstruction error?

For the enthusiast. Train and test an SVM on the digit classification problem of Week 5, now using the PCA features extracted from the images. How does the classification error vary as the number of PCA features is increased from 1 to the original dimension of the data.

Prepare a short report with diagrams illustrating the application of PCA on the digit images.