Homework 6

Study Chapter 12 from the Classnotes and the slides on the class website.
Read Chapter 12.1 and 12.3 from Bishops book.

1) Derive a linear transformation $y=Ax$ based on principal components which has the property that the covariance $\text{Cov}(y)=\text{Identity}$.

2) Imagine you have a collection of gray-level images of faces. Provide pseudo-code to compute the “eigenfaces” of this collection. (see below for 4 examples of eigenfaces).

![Eigenfaces Example](image)

3) Perform a PCA on the Experian data. Make sure you first center the data. Extract the first two eigenvalues and eigenvectors.

4) Project the data down to a two-dimensional subspace and produce a scatter plot of the data. Make sure you plot each of the three classes differently (using color or different markers).

5) Define the following kernel:

$$K(x,y) = \exp\left[ -\frac{\lambda}{2} \sum_{k=1}^{K} \| x_k - y_k \|^2 \right].$$

Compute the kernel matrix (or Gram matrix) on the first 100 data-cases of the Experian dataset (basically as many as you can handle computationally). Make sure you first center the data to have zero mean and scale the data to have unit variance for all attributes separately.

d) Project the data down to a two dimensional subspace using the features computed from kernel-PCA. Can you see any differences between PCA and kernel-PCA?