Master of bioinformatics

Support Vector Machines

Assignment 3: Unsupervised Learning

Spring 2016

Author:
Cedric Lood

Supervisors:
Dr. Carlos Alaiz
Dr. Emanuele Frandi
Prof. Johan Suykens

May 23, 2016

Contents

Context

The analysis presented in this report was produced for the class of "Support Vector Machines: methods and applications" at KU Leuven (Spring 2016). The goal is to display understanding of the principles behind support vector machines and of how to work out good solutions using these techniques. This third report focuses on unsupervised learning (kernel PCA) using Least-Squares SVM (LS-SVM). The implementation was done using the MatLab environment (v2015a) and the libraries for LS-SVM developed at KU Leuven [1].

1 Kernel Principal Component Analysis

In this section, we explore the synthetic dataset illustrated on figure 1 to explore the relationship between the choices of kernel, hyper-parameters and the number of components.

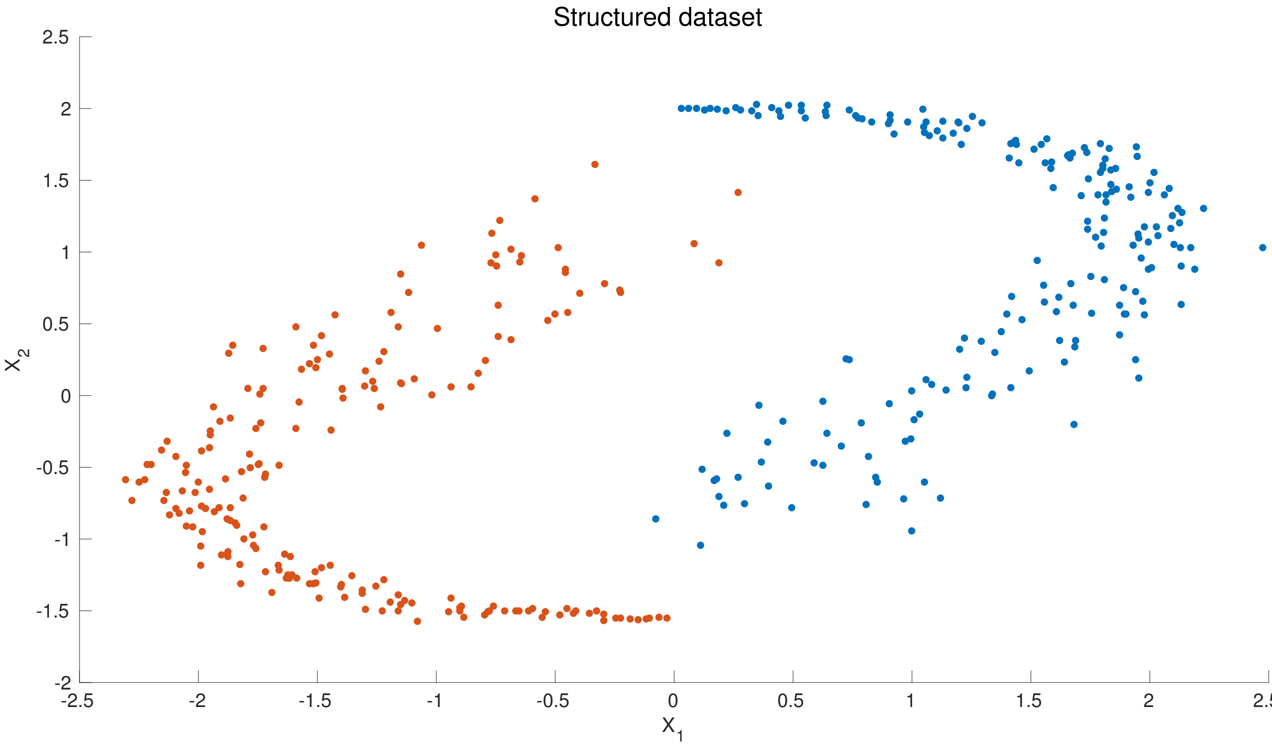

Figure 1: Synthetic dataset Yin-Yang

Figure 2 allows us to visualize what happens when you increase the number of component in a given setting (here $\sigma^2 = 0.4$). We can see that the denoising works really well once the number of components has increased to 4 and above. Figure 3 illustrates the impact of σ^2 on the denoising for a number of components fixed to 6.

Given the structure of the dataset, linear PCA gives very poor results in terms of denoising (see figure 4). The maximum amount of principal component that can be obtained using linear PCA corresponds to the dimensionality of the original input space. For the kernel PCA, since we are working in the feature space, the dimensionality can be higher, and corresponds to the dimensionality of the kernel matrix.

[1] http://www.esat.kuleuven.be/sista/lssvmlab/

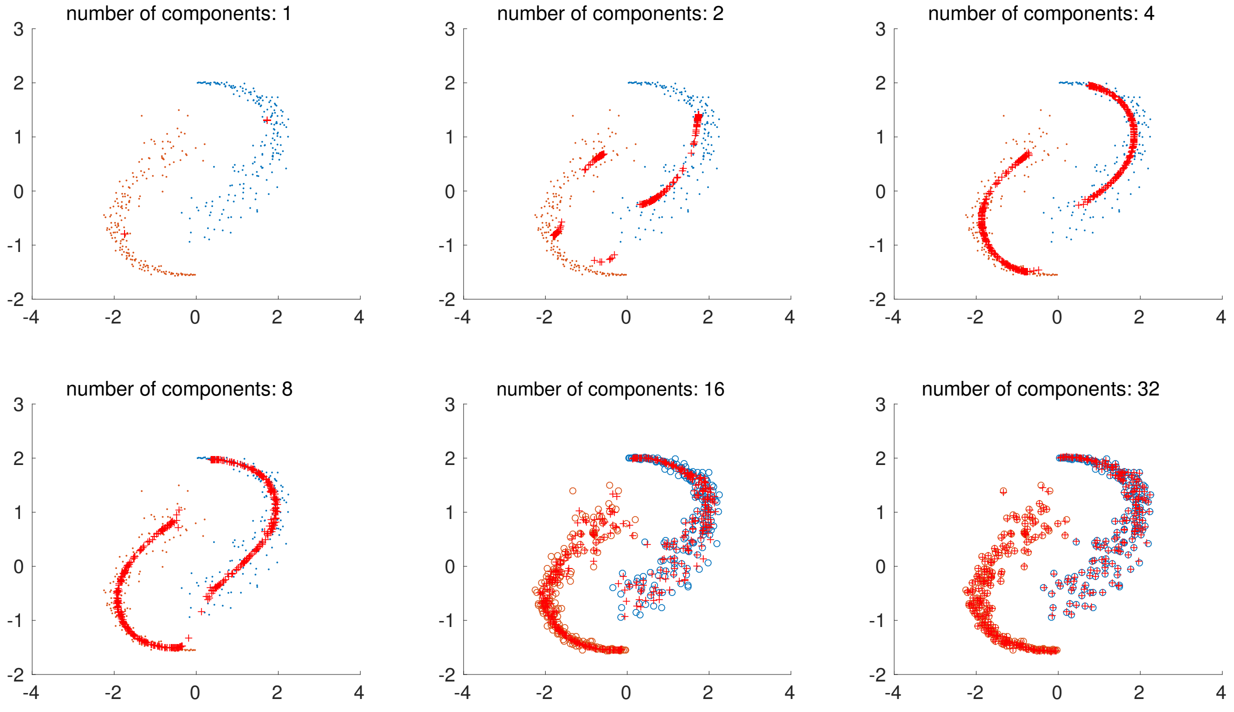

Figure 2: Denoising for various values of components

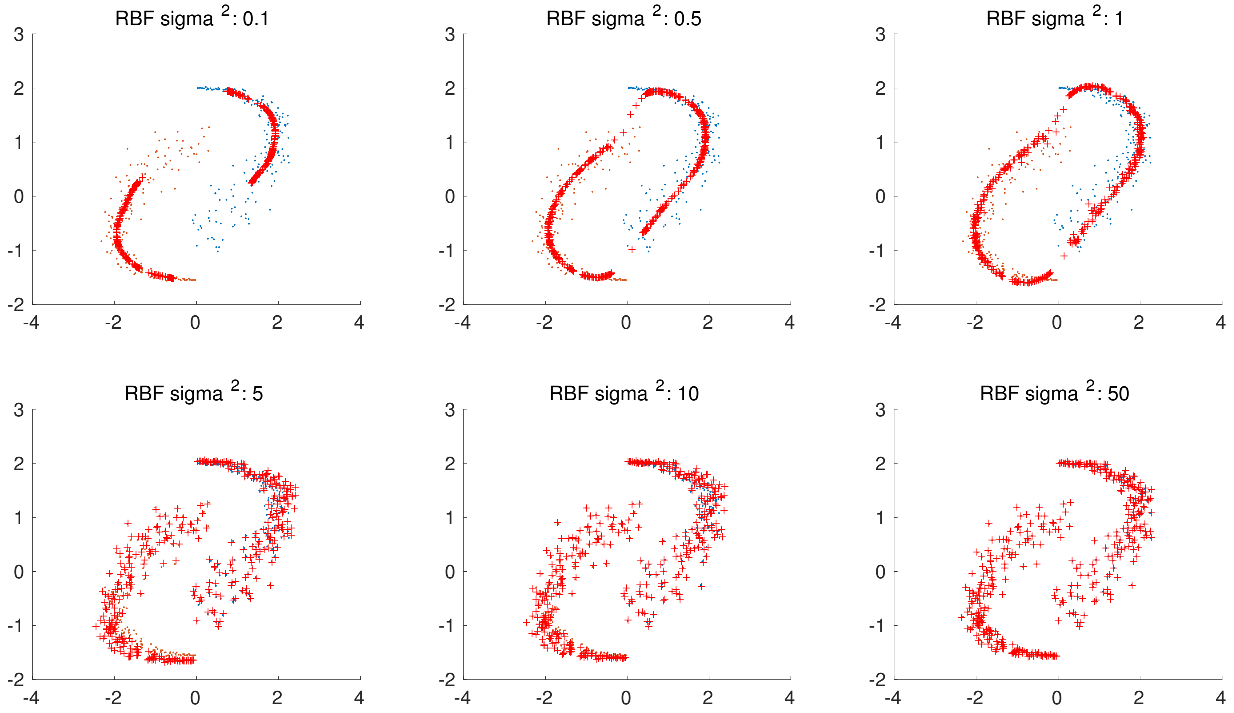

Figure 3: Denoising for various values of σ^2

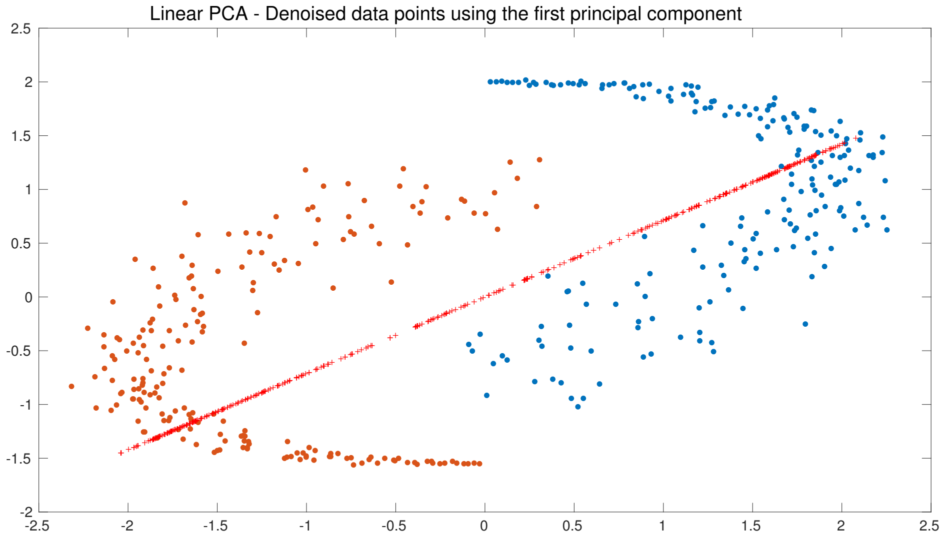

Figure 4: Denoising in linear PCA setting

2 Handwritten Digit Denoising

The results of the denoising process are reported on figures **??** and **??**. As can be observed, the performance of the kernel PCA in terms of denoising are much better. Basically, the linear PCA decomposition into 256 components, followed by the reconstruction using progressively more and more components (256 in the last row) returns the original noisy image. Where as the kernel PCA is able to extract out the gaussian noise during the reconstruction, returning a very convincing, denoised image which is similar to the original noiseless one.

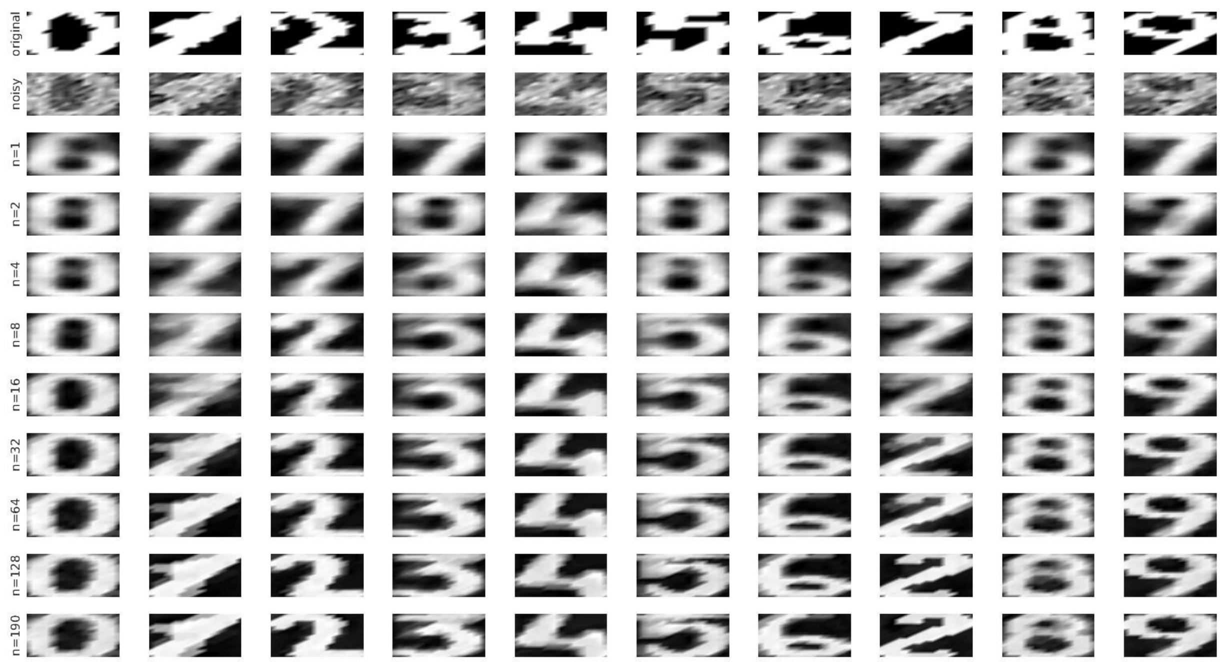

Figure 5: Denoising digits using a kernel PCA approach

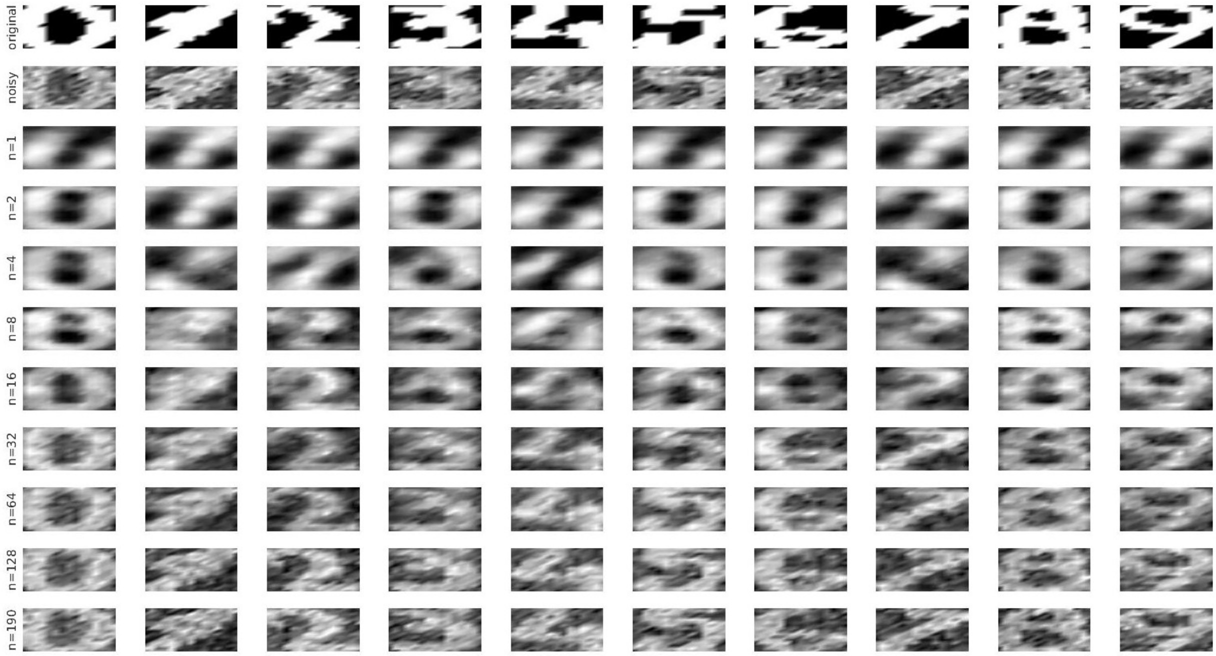

Figure 6: Denoising digits using a linear PCA approach

3 Spectral Clustering

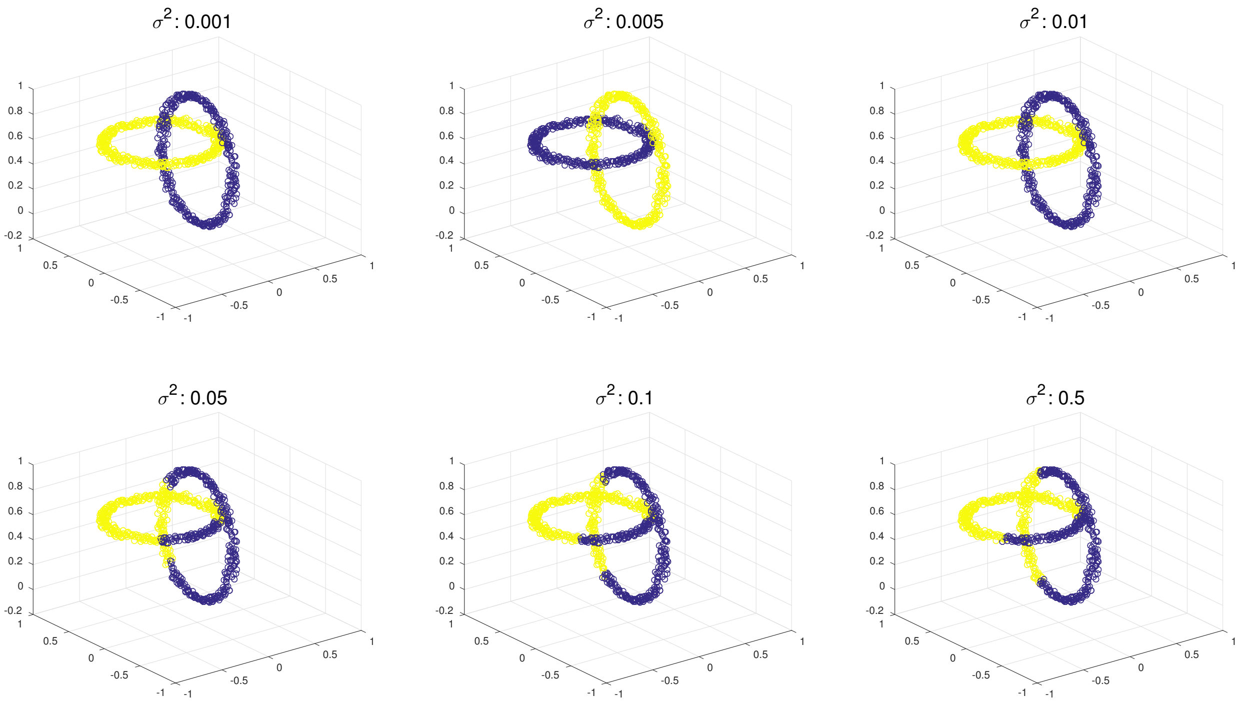

Figure 7: Clustering 1

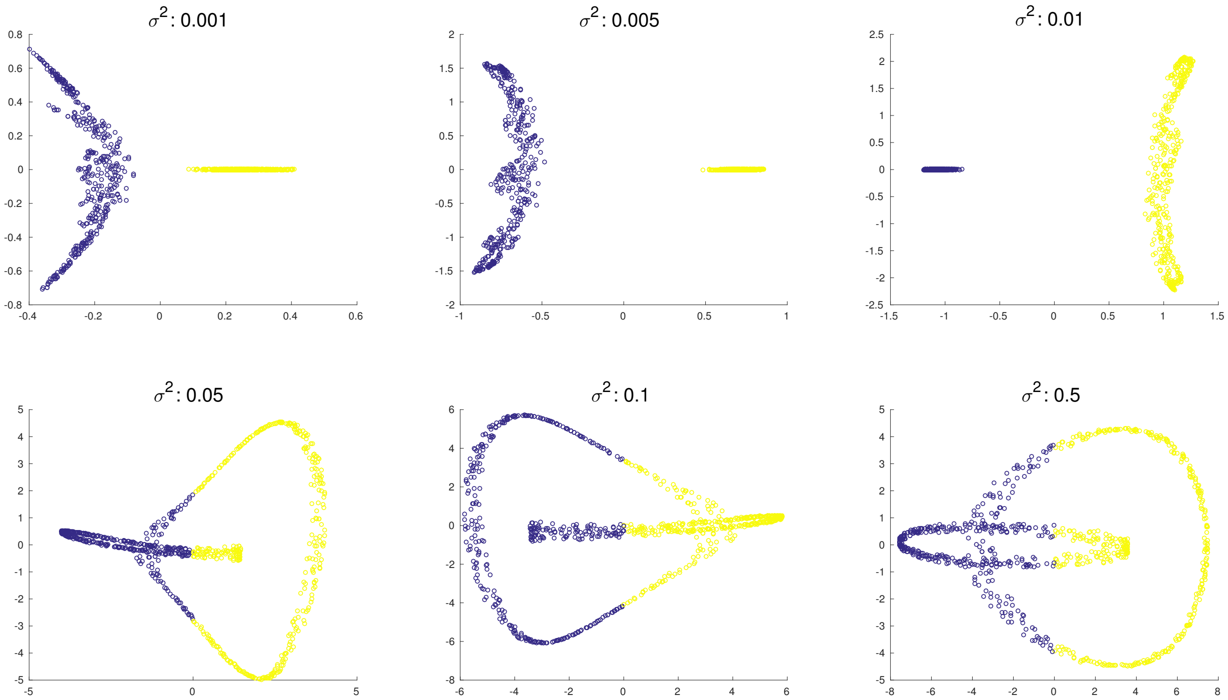

Figure 8: Clustering 2

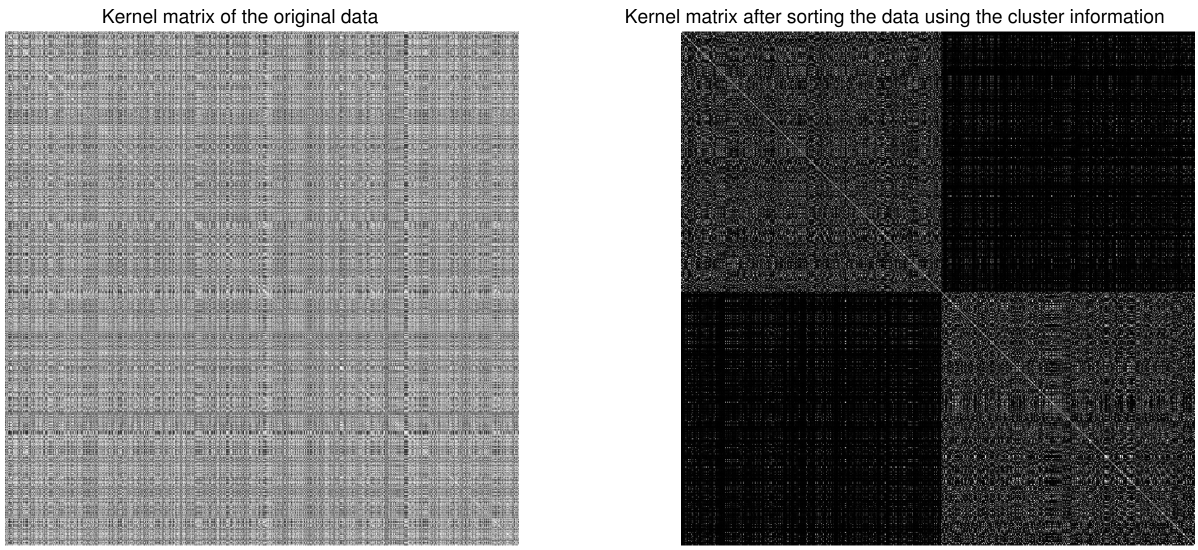

Figure 9: Matrix

4 Fixed-size LS-SVM

5 Applications

5.1 Handwritten Digit Denoising

5.2 Shuttle (statlog)

5.3 California