Ludwig-Maximilians-Universität München
Institut für Informatik
Prof. Dr. Peer Kröger
Yifeng Lu

Knowledge Discovery in Databases II
SS 2018

Exercise 4: Dimentionality Reduction

Exercise 4-1 Mutual Information

Calculate the mutual information of the following dataset (from Exercise 3-1).

Grade	took part	enjoyed summer
1	10	5
2	25	8
3	35	7
4	20	16
5	12	48

Exercise 4-2 SVD

Given matrix $A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$:

(a) Find the singular values of the matrix A.

(b) Energy is defined in terms of the singular values. If we reduce the dimensionality of matrix A to 1 using SVD, how many percentage of energy can be retained?

Exercise 4-3 Kernel PCA

Download the 2-dimensional dataset *kpca.csv*. Many programming languages provide the `PCA()` function, such as R or python. Choose the one you preferred. Design your kernel functions to achieve linear splitting and compare the results of PCA with and without a kernel function.
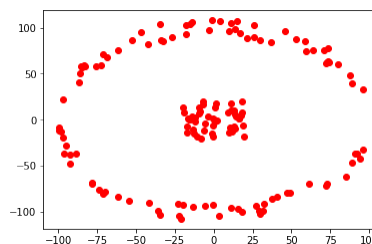

Abbildung 1: *kpca.csv* plot