

EXERCISE SESSION Spectral Clustering: ADVANCED MACHINE LEARNING
COURSE – EPFL – Lecturer A. Billard

A key step in spectral clustering is the assumption that the non-zero entries on the eigenvectors with small eigenvalues code for groups of points that are close to another. In this exercise, we will get an intuition of why this assumption may be correct.

- 1) Consider a two-dimensional dataset composed of two points.
 - a. Build a similarity matrix using a threshold function on Euclidian (norm-2) distance. The metric outputs 1 if the points are close enough according to a threshold and zero otherwise. Consider two cases: when the two datapoints are close or far.
 - b. Build the Laplacian in each case and discuss the eigenvalues and eigenvectors.
- 2) Redo (a) using this time a RBF kernel to build the similarity matrix.
- 3) Redo (a) using a homogeneous polynomial kernel to build the similarity matrix.