

ML 4/M

Lab exercise: K-means clustering

Simon Rogers, March 2017

Aims and outline

To implement the K -means clustering algorithm.

Tasks

- Download the skeleton notebook (`K_means_skeleton.ipynb`)
- The first couple of cells generate a dataset with clear cluster structure
- The data is stored in a 60×2 matrix \mathbf{X}
- After the data is plotted, the various things needed by K-means are created: the number of clusters is stored in a variable K , and \mathbf{z} is defined as an $N \times K$ matrix that will store the current cluster memberships of the N points ($z_{nk} = 1$ if object n is in class k).
- Your task is to write the next cell which should run the K-means iterations. Comments have been provided to help.
- Once written, try experimenting with different values of K , or change the data to see how the algorithm behaves.