Exercise 20 - PCA

This exercise gives you hands-on experience about PCA.

- **(3 Points) PCA** Complete the function `PCA_diy.m` where you input the raw data `X` and specify the number of reduced dimension `d` and output the projection on the reduced dimension space.

- **(3 Points) Kernel PCA** Complete the function `kPCA_diy.m` with Gaussian kernel with the same input and output format as in `PCA_diy.m`.

- **(2 Points)** Now apply your code (PCA & kPCA) to the provided dataset `pca_data.mat`, visualize your results by running the script `ex20_pca.m` and discuss the results.

Hints:

- For kernel PCA, you can find more information from `http://www.eecs.berkeley.edu/~wainwrig/stat241b/scholkopf_kernel.pdf`

Exercise 21 - Semi-supervised Learning

a. **(4 Points)** Prove the following equality: Let $G = (V, E)$ be an undirected graph with vertex set $V = \{v_1, \ldots, v_n\}$. The weighted adjacency matrix of the graph is the matrix $W = (w_{ij})_i, j = 1, \ldots, n$. The degree matrix D is defined as the diagonal matrix with the degrees $d_1, \ldots, d_n$ on the diagonal. For every vector $f \in \mathbb{R}^n$ we have,

$$\Big\langle f, (\mathbb{1} - D^{-\frac{1}{2}} W D^{-\frac{1}{2}})f \Big\rangle_{\mathbb{R}^n} = \frac{1}{2} \sum_{i,j=1}^{n} w_{ij} \Big(\frac{f_i}{\sqrt{d_i}} - \frac{f_j}{\sqrt{d_j}} \Big)^2.$$

Submission instructions

- We accept both handwritten and electronic submissions. So you can choose what is more convenient for you. In any case, you should specify full names and immatriculation IDs of all team members. Obviously, programming tasks you can submit only electronically.

- Handwritten submissions should be submitted in the lecture hall of Monday's lecture (before the lecture starts).

- Electronic submissions should be zipped, containing the m-files (`Basis` etc.), your plots (png files) and the matlab data files (.mat) and emailed to the corresponding tutor:

  a. Apratim Bhattacharyya (Wednesday 8-10): abhattac@mpi-inf.mpg.de
  b. Maksym Andriushchenko (Thursday 8-10): s8mmandr@stud.uni-saarland.de

    c. Max Losch (Friday 16-18): mlosch@mpi-inf.mpg.de

If not all 3 students belong to the same tutorial group, then you should email your submission to **only** one tutor (e.g. to the tutor of the first author of your homework), so please do not put other tutors in copy of the email.

The email subject must have the following form: "[ML18/19 Exercise] Sheet X", where X is the number of the current exercise sheet. Then please specify in the email full names and immatriculation IDs of all team members. Then please attach all your files as a single zip archive, which consists of your immatriculation IDs, e.g. "2561234_2561235_2561236.zip".

- Reminder: you should submit in groups of 3. Otherwise, we will later on merge the groups smaller than 3 students.