Queen Mary University of London

Main Examination period 2023 – May/June – Semester B

MTH793P: Advanced machine learning

Duration: 4 hours

The exam is available for a period of **4 hours**, within which you must complete the assessment and submit your work. **Only one attempt is allowed – once you have submitted your work, it is final**.

All work should be **handwritten** and should **include your student number**.

You should attempt ALL questions. Marks available are shown next to the questions.

In completing this assessment:

- **You may use books and notes.**
- **You may use calculators and computers, but you must show your working for any calculations you do.**
- **You may use the Internet as a resource, but not to ask for the solution to an exam question or to copy any solution you find.**
- **You must not seek or obtain help from anyone else.**

When you have finished:

- scan your work, convert it to a **single PDF file**, and submit this file using the tool below the link to the exam;
- e-mail a copy to **maths@qmul.ac.uk** with your student number and the module code in the subject line;
- with your e-mail, include a photograph of the first page of your work together with either yourself or your student ID card.

Examiners: 1st **Dr. N. Perra,** 2nd **Dr. N. Otter**

Question 1 [20 marks].

You are given the three data points coloured in blue in Figure 1. You are tasked to add two points, the red points in Figure 1, by interpolation. Considering the problem as a semi-supervised prediction task, complete the vector $\mathbf{y} = (1, ?, 3, ?, 2)$ which represents the y-coordinates of blue (known) and red (unknown) points.
In particular:

(a) By ordering nodes from left to right, write down the incidence matrix $\mathbf{M}$, and using the incidence matrix write down the Laplacian matrix $\mathbf{L}$. **[5]**

(b) We know that the Laplacian matrix can be expressed as $\mathbf{L} = \mathbf{D} - \mathbf{A}$ where $\mathbf{D}$ is a diagonal matrix whose elements are the degree of each node (i.e., data point) and $\mathbf{A}$ is the adjacency matrix. Write down the expressions for these two matrices and verify that $\mathbf{L} = \mathbf{D} - \mathbf{A}$. In doing so, keep the order of nodes as before: from left to right. **[5]**

(c) We know that this problem leads to a normal equation of the form

$$P_{I_1/I_2}^{\top} \mathbf{L} P_{I_1/I_2} \hat{\mathbf{w}} = -P_{I_1/I_2}^{\top} \mathbf{L} P_{I_2} \mathbf{v}$$

where $\mathbf{v}$ is the vector of known y-values. Write down the expressions for P_{I_2} and P_{I_1/I_2}. **[5]**

(d) Find $\hat{\mathbf{w}}$ by solving the normal equation. Show all the steps and do not just infer/read the values from the figure. **[5]**

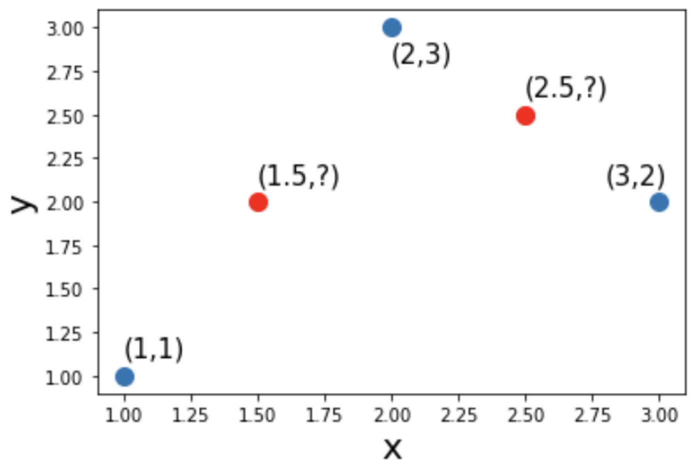

Figure 1: Known points are shown in blue. The red points describe new points we need to add by interpolation.

Question 2 [25 marks].

Consider the five data points plotted in Figure 2. Assume that they are ordered clockwise such that $\mathbf{p_1} = (1, 1)^\top$, $\mathbf{p_2} = (2, 1)^\top$, $\mathbf{p_3} = (-1, -1)^\top$, $\mathbf{p_4} = (-2, -1)^\top$, $\mathbf{p_5} = (-1, 1)^\top$.

(a) Write the pairwise distance matrix $\mathbf{d}$ whose elements d_{ij} describe the Euclidian distance between point $\mathbf{p_i}$ and $\mathbf{p_j}$. **[10]**

(b) Starting from the distance matrix $\mathbf{d}$ let us build a graph $\mathbf{G}$ such that:

- data points are connected only if $d_{ij} \leqslant 2$,
- the weight of each pair of connected nodes is defined as $w_{ij} = d_{ij}^{-1}$,

draw the graph, write down the adjacency matrix $\mathbf{A}$, the diagonal matrix $\mathbf{D}$, and the Laplacian $\mathbf{L}$ of the correspondent graph. **[5]**

(c) just by looking at the Laplacian matrix what can we say about i) the value of the its second smallest eigenvalue λ_2, ii) the graph $\mathbf{G}$. **[5]**

(d) Consider a connected graph $\mathbf{G}(N, E)$ characterised by the adjacency matrix $\mathbf{A} \in \mathbb{R}^{N \times N}$, the diagonal degree matrix $\mathbf{D} \in \mathbb{R}^{N \times N}$ and the Laplacian matrix $\mathbf{L} \in \mathbb{R}^{N \times N}$. Consider a vector $\mathbf{p} \in \mathbb{R}^{N \times 1}$ that satisfies the following equation $\mathbf{p} = \mathbf{A}\mathbf{D}^{-1}\mathbf{p}$. Prove that $\mathbf{p} = c\mathbf{D}\mathbb{1}$, where c is a constant and $\mathbb{1} \in \mathbb{R}^{N \times 1}$ is a vector whose components are all equal to one. **Hint**: Use the properties of the Laplacian matrix. **[5]**

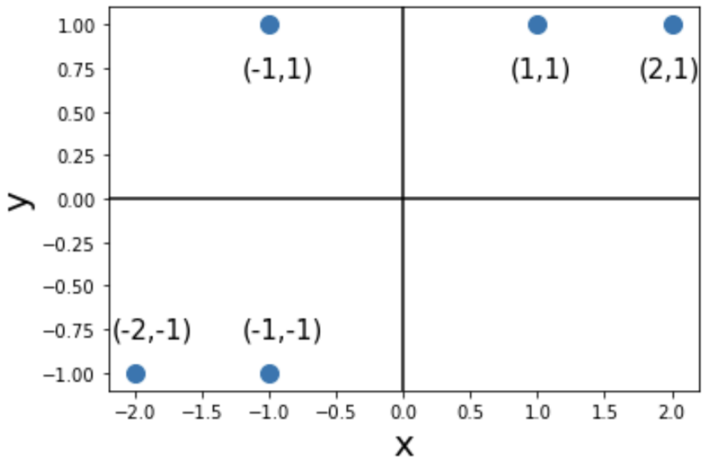

Figure 2: Data points

Question 3 [35 marks].

Consider the following six data points: $\mathbf{p_1} = (1,1)^\top$, $\mathbf{p_2} = (2,2)^\top$, $\mathbf{p_3} = (1,-1)^\top$, $\mathbf{p_4} = (-1,-1)^\top$, $\mathbf{p_5} = (-2,-2)^\top$, $\mathbf{p_6} = (-1,1)^\top$.

(a) Write down the correspondent $\mathbf{X} \in \mathbb{R}^{2\times 5}$ matrix and show that the data is already centred. **[5]**

(b) Find the principal components of $\mathbf{X}$. **[10]**

(c) Compute the projections $y_1, \ldots, y_6 \in \mathbb{R}^1$ of $\mathbf{p_1}, \ldots, \mathbf{p_6} \in \mathbb{R}^2$ on the first principal component. **[5]**

(d) Compute the reconstructions of $\mathbf{p_1}, \ldots, \mathbf{p_6}$ using the first principal components, denoted $\hat{\mathbf{p}}_1, \ldots, \hat{\mathbf{p}}_6 \in \mathbb{R}^2$. **[5]**

(e) In a 2D axis system, plot the following:

- The original points $x_1, \ldots, x_6$,
- The reconstructed points $\hat{x}_1, \ldots, \hat{x}_6 \in \mathbb{R}^2$,
- The principal components (directions).

[5]

(f) Consider the following matrix:

$$\mathbf{M} = \begin{pmatrix} 3 & 6 & 1 & 9 \\ 1 & 2 & \frac{1}{3} & 10+d \\ 2 & 14+d & \frac{2}{3} & 6 \end{pmatrix}$$

where d is the last digit of your student ID number. Find the decomposition $\mathbf{M} = \mathbf{L} + \mathbf{E}$ where $\mathbf{E}$ is a sparse matrix (with at most 3 nonzero entries), and $\mathbf{L}$ is a low-rank matrix (lowest rank possible). **[5]**

Question 4 [20 marks].

(a) Compute the nuclear norm and the rank of the matrix:

$$\mathbf{M} = \begin{pmatrix} 1 & 2 & 1 \\ 1 & 2 & 0 \end{pmatrix}$$

[10]

(b) Consider a matrix $\mathbf{M}$, show that $\|M\|_F^2 = \sum_{i=1}^{r} \sigma_i^2$ where the σ_i are the singular values. **[5]**

(c) Consider a matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$ and its singular value decomposition $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$. Prove that $c\mathbf{U}^\top\mathbf{X}\mathbf{V}\mathbf{1} = (c\sigma_1, \ldots, c\sigma_r)^\top$ where $\mathbf{1} \in \mathbb{R}^{n \times 1}$ is a constant vector with all components equal to one, and $c \in \mathbb{R}$. **[5]**

End of Paper.