Late-Summer Examination period 2023

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 [25 marks].

Consider a graph $\mathbf{G}(V, E)$ defined by the following list of undirected edges $e_1 = (1, 2)$, $e_2 = (2, 3)$, $e_3 = (1, 3)$, $e_4 = (3, 4)$, $e_5 = (4, 5)$, $e_6 = (5, 6)$, $e_7 = (4, 6)$. The edges are weighted as follows $w_{1,2} = 10$, $w_{2,3} = 10$, $w_{1,3} = 5$, $w_{3,4} = 2$, $w_{4,5} = 20$, $w_{5,6} = 15$, $w_{4,6} = 10$

(a) Draw the graph and write down the incidence matrix $\mathbf{M}$. **[5]**

(b) Using the incidence matrix write down the Laplacian matrix $\mathbf{L}$. **[5]**

(c) Write down the adjacency matrix $\mathbf{A}$ and the matrix $\mathbf{D}$ whose diagonal elements are the weighted degree (i.e., strength) of each node. Verify that $\mathbf{L} = \mathbf{D} - \mathbf{A}$. **[5]**

(d) Assume that we gather partial information about some category c of nodes 1 and 4, namely $c_1 = 1$ and $c_4 = 0$. Using the network information in a semi-supervised setting, you are tasked to predict the category of the other four nodes. We know that this problem leads to the following equation:

$$P_{I_1/I_2}^{\top} \mathbf{L} P_{I_1/I_2} \hat{\mathbf{w}} = -P_{I_1/I_2}^{\top} \mathbf{L} P_{I_2} \mathbf{v} \tag{1}$$

where $\mathbf{v}$ is the vector of known categories.

- Write down the expressions for P_{I_2} and P_{I_1/I_2}, **[5]**
- Find $\hat{\mathbf{w}}$ by solving the normal equation. **[5]**

 Continue to next page

Question 2 [40 marks].

(a) Considering the following data points $p_1 = -3$, $p_2 = -1$, $p_3 = 1$, $p_4 = 2$, $p_5 = 3$, $p_6 = 5$, $p_7 = 8$, cluster them applying k-means starting with centroids $\mu_1^{(0)} = 0$ and $\mu_2^{(0)} = 9$. Write out all steps of the algorithm by hand. **[10]**

(b) Compute the Rand Index considering as $\mathcal{P}_1$ the partition outcome of the clustering in the previous point and $\mathcal{P}_2 = (C'_1, C'_2, C'_3)$ where $C'_1 = (p_1, p_2)$, $C'_2 = (p_3, p_4)$ and $C'_3 = (p_5, p_6, p_7)$. Write out all steps by hand. **[15]**

(c) We are now given a representation of the points in a two dimensional space $\mathbf{p}_1 = (-1,-1)^\top$, $\mathbf{p}_2 = (-3,-2)^\top$, $\mathbf{p}_3 = (1,0)^\top$, $\mathbf{p}_4 = (2,0)^\top$, $\mathbf{p}_5 = (3,4)^\top$, $\mathbf{p}_6 = (5,2)^\top$, $\mathbf{p}_7 = (8,2)^\top$. Cluster the data points applying k-means starting with centroids $\mu_1^{(0)} = (0,-1)^\top$ and $\mu_2^{(0)} = (1,1)^\top$. Write out all steps of the algorithm by hand. **[15]**

Question 3 [20 marks]. Consider the following data points $\mathbf{p}_1 = (1, 2)^\top$, $\mathbf{p}_2 = (2, 3)^\top$, $\mathbf{p}_3 = (4, 1)^\top$.

(a) Write down the correspondent $\mathbf{X} \in \mathbb{R}^{2\times 3}$ data matrix, compute its singular values, and the left singular vectors $\mathbf{X}$. **[10]**

(b) Compute the matrix $\text{soft}_\tau(\mathbf{\Sigma})$ obtained by applying the soft thresholding operator to each element of the matrix of singular values $\mathbf{\Sigma}$. Set τ equal to the last digit of your student ID. Compute the nuclear norm of $D_\tau(\mathbf{X}) = \mathbf{U}\text{soft}_\tau(\mathbf{\Sigma})\mathbf{V}^\top$ and compare it with the nuclear norm of the original matrix ($\mathbf{U}$ and $\mathbf{V}$ are the left and right singular vectors respectively). **[5]**

(c) Compute a lower rank approximation $\hat{\mathbf{L}} \in \mathbb{R}^{2\times 3}$ of the matrix $\mathbf{X}$ by hand, considering $rank(\hat{\mathbf{L}}) = 1$ and such that $\|\hat{\mathbf{L}} - \mathbf{X}\| \leqslant \|\mathbf{L} - \mathbf{X}\|$ for all $\mathbf{L} \in \mathbb{R}^{2\times 3}$ and $rank(\mathbf{L}) = 1$. **[5]**

 Continue to next page

Question 4 [15 marks].

(a) Consider the matrix:

$$\mathbf{X} = \begin{pmatrix} 1 & 5 & 0 \\ 0 & 1 & 1 \end{pmatrix} \tag{2}$$

diagonalise the matrix $\mathbf{X}\mathbf{X}^\top$ and discuss its connection with the SVD of $\mathbf{X}$. **[10]**

(b) Consider a matrix $\mathbf{M} \in \mathbb{R}^{m \times n}$. Show that $Trace(\mathbf{M}\mathbf{M}^\top + \mathbf{M}^\top\mathbf{M}) = 2\sum_{i=1}^{r} \sigma_i^2$ where σ_i are its singular values. **[5]**

End of Paper.