Queen Mary
University of London

Late-Summer Examination period 2023

MTH786P: Machine learning with Python

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/2nd Examiners: N. Perra, N. Otter

Question 1 [45 marks].

(a) Consider the following matrix $\mathbf{A} = \begin{pmatrix} a & 1 \\ 1 & -1 \end{pmatrix}$ where $a \in \mathbb{R}$. Compute the eigenvalues and correspondent eigenvectors of $\mathbf{A}$. Discuss for which values of a the eigenvalues are real. **[15]**

(b) Consider the following matrix $\mathbf{M} = \begin{pmatrix} 2 & 2 \\ -1 & 1 \\ 0 & 0 \end{pmatrix}$ Compute the singular values, left and right singular vectors. **[15]**

(c) Consider a matrix $\mathbf{B} \in \mathbb{R}^{n,n}$, a matrix $\mathbf{C} \in \mathbb{R}^{n,n}$, a vector $\mathbf{x} \in \mathbb{R}^{n,1}$, a vector $\mathbf{v} \in \mathbb{R}^{n,1}$ and the function $E(\mathbf{x}) = \frac{1}{3}\|\mathbf{B}(\mathbf{x}+\mathbf{v}) + \mathbf{C}(\mathbf{x}-\mathbf{v})\|^2$. Compute the gradient of the energy function i.e., $\nabla E(\mathbf{x})$ **[15]**

Question 2 [25 marks].

Consider the following data samples $\left(x^{(1)}, y^{(1)}\right) = (1, 2)$, $\left(x^{(2)}, y^{(2)}\right) = (0, 1)$, $\left(x^{(3)}, y^{(3)}\right) = (2, 3)$

(a) Write down, in explicit matricial form, the normal equation assuming a ridge regression. **[5]**

(b) Determine the solution of the normal equation assuming a ridge regression. **[5]**

(c) Let us now assume that you made some errors measuring the output variables $y^{(i)}$ with $i \in \{1, 2, 3\}$. The perturbed measurements $\mathbf{y}_\delta$ read $y_\delta^{(1)} = 2 + \epsilon$, $y_\delta^{(2)} = 1$ and $y_\delta^{(3)} = 3 - \epsilon$. Determine the solution of the normal equation considering these perturbed samples and considering the same initial data matrix. **[5]**

(d) Compute the error between $\hat{\mathbf{w}}$ and $\hat{\mathbf{w}}_\delta$ in the Euclidean norm. **[5]**

(e) Compute the data error $\delta := \|\mathbf{y} - \mathbf{y}_\delta\|$. **[5]**

Question 3 [30 marks].

Consider the following data samples $\left(x^{(1)}, y^{(1)}\right) = (-1, -1)$, $\left(x^{(2)}, y^{(2)}\right) = (0, 1)$, $\left(x^{(3)}, y^{(3)}\right) = (1, 2)$, $\left(x^{(4)}, y^{(4)}\right) = (2, 3)$

(a) Compute the MSE for a 1-parameter model by hand:

$$MSE(w^{(0)}) = \frac{1}{2s} \sum_{i=1}^{s} |y^{(i)} - w^{(0)}|^2$$

considering $w^{(0)} \in \{1, 2, 3\}$. Which of the three values minimizes the MSE? **[10]**

(b) a new data sample is added $\left(x^{(5)}, y^{(5)}\right) = (3, 1)$. Evaluate new error measure and corresponding minimiser considering the same three possible values for $w^{(0)}$ **[5]**

(c) Repeat the same exercise for what is known as the Mean Absolute Error (MAE)

$$MSE(w^{(0)}) = \frac{1}{2s} \sum_{i=1}^{s} |y^{(i)} - w^{(0)}|$$

What do you observe, in particular with regards to the outlier $y^{(5)}$? **[10]**

(d) Discuss the convex properties of the function $f(x) = a \sin(x)$ considering $0 \leqslant x \leqslant 2\pi$ and $a \in \mathbb{R}$. **[5]**

End of Paper.