Two hours

EXAM PAPER MUST NOT BE REMOVED FROM
THE EXAM ROOM AND MUST BE RETURNED

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Modelling and visualisation of high-dimensional data

Date: Tuesday 26th January 2016

Time: 14:00 - 16:00

Answer ALL Questions in Section A
Write your answers directly on the exam paper. Only answers written in the boxes on the exam paper will be marked.

Answer ALL Questions in Section B, use a separate answerbook for this Section

---

This is a CLOSED book examination

The use of electronic calculators is permitted provided they are not programmable and do not store text

Section A

This Section contains Multiple Choice Questions and is therefore restricted

Section B

Answer *all* questions from this section.

1. The *Locally Linear Embedding* (LLE) is a typical non-linear manifold learning algorithm. Give all the cost functions used to derive the LLE algorithm and explain their roles in terms of manifold learning and state the hyper-parameter that may considerably affect the performance. (5 marks)

2. Face recognition has been widely used for personal identity authentication. The use of salient features of facial images often significantly improves the performance of a face recognition system in terms of both recognition accuracy and computational efficiency. You are asked to apply the *linear discriminative analysis* (LDA) learnt from this course unit to generate a facial representation (a facial feature set) that effectively captures salient facial features and apply it to build up a face recognition system. Describe all the main steps required in establishing the face recognition system with your chosen techniques and discuss the limitation of LDA in this application. It is essential to give the algorithmic details on how to extract a facial representation via learning and how to use the facial representation during the authentication. (10 marks)

3. *Principal Component Analysis* (PCA) is an effective and popular dimension reduction method. Making use of the derivation in obtaining the top principal component learnt from the course unit, prove that, in PCA, the linear projection onto an M-dimensional subspace that maximises the variance of the projected data is defined by the M eigenvectors of the covariance matrix of an N-dimensional data set corresponding to the M largest eigenvalues ($N > M$). (10 marks)

END OF EXAMINATION