SCHOOL OF ENGINEERING

ADVANCED CONCEPTS IN SIGNAL PROCESSING

PGEE11020

Exam Date: **16/05/2014** From and To: **1430-1630** Exam Diet: **May 2014**

Please read full instructions before commencing writing

Exam paper information

- Paper consists of 2 Sections
- Candidates to answer THREE questions
- Section A: (One question) Answer whole Section
- Section B: Answer TWO out of THREE questions

Special instructions

- Students should assume reasonable values for any data not given in a question nor available on a datasheet, and should make any such assumptions clear on their script.
- Students in any doubt as to the interpretation of the wording of a question, should make their own decision, and should state it clearly on their script.
- Please write your name in the space indicated at the right hand side on the front cover of the answer book. Also enter you examination number in the appropriate space on the front cover.
- Write **ONLY** your examination number on any extra sheets or worksheets used and firmly attach these to the answer book(s).
- This examination will be marked anonymously.

Special items

- None

Convenor of Board of Examiners: **Professor B Mulgrew**
External Examiner: **Professor D Bull**

SECTION A

a) Explain the Newton method and gradient descent method for parameter optimization. Your answer should explain the relationship between the two methods and indicate the strengths and weaknesses of each. **(5)**

b) Suppose that the number of hidden units in a neural network is to be chosen automatically to achieve the best classification. Discuss why it would be wrong simply to use error of the training set for evaluation. **(2)**

c) Describe the technique of cross-validation giving an example of how you might use it to solve the problem in part b). **(3)**

d) Determine whether the means $\mu_1 = \binom{1}{1}$ and $\mu_2 = \binom{5}{0}$ form a possible solution to the k-means clustering of the following 2-dimensional data:

$$(0,0), (5,-1), (0,2), (5,1), (3,1).$$

(5)

e) Explain the operation of the Hidden Markov Model (HMM) defined by the transition matrix, **A**, and the observation matrix, **B**, given by:

$$\mathbf{A} = [a_{ij}] = \begin{pmatrix} 0.8 & 0.1 & 0.1 \\ 0 & 0.5 & 0.5 \\ 0 & 0 & 1 \end{pmatrix}, \mathbf{B} = [b_{jk}] = \begin{pmatrix} 0.1 & 0.9 & 0 \\ 0.8 & 0.2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

where the indices i, j and k range from 1 to 3. Your answer should define the state transition diagram and clearly indicate the roles of the states and the observed symbols. **(5)**

SECTION B

Question B1

a) Define the perceptron criterion for a binary linear classification task. Hence derive the batch perceptron algorithm, identifying any of the important terms. **(6)**

b) Under what condition do we expect the batch perceptron algorithm to converge? **(1)**

c) Consider the following data points $(x_1, x_2)_i$, taken from two categories, ω_1 and ω_2:

$$\omega_1 : (1,4)(3,2)$$

$$\omega_2 : (0,3)(1,2)$$

Given an initial estimate of $w_0 = -1, w_1 = 0, w_2 = 0$ calculate the estimate for the new weights after two iterations of the batch perceptron algorithm with a step size of 1.0. **(6)**

d) Define the *support vectors* of a linear classifier and explain their importance. **(2)**

e) Sketch the data from part c) indicating the decision boundary with the maximum margin. Hence determine the weight vector for a maximum margin classifier. **(4)**

f) Identify the support vectors for the maximum margin classifier in part f). **(1)**

Question B2

a) Explain what is meant by the term *Maximum Likelihood* and discuss its role in machine learning problems. **(4)**

b) Suppose that we are given a set of data $\{x_i\}$ with class labels, $\{t_i\}$, of which the fraction α are labelled as belonging to class ω_1. Write down the likelihood function for the parameter p that estimates the prior probability of class ω_1: $p = \hat{P}(\omega_1)$. **(2)**

c) Hence calculate the maximum likelihood solution for p. **(3)**

d) Write down the formula for the Bayes Decision Boundary for a two category multi-dimensional classification problem with Gaussian class conditional probability distributions. **(3)**

e) Given the data in figure B2a drawn from two Gaussian distributions, calculate the Bayes decision boundary stating any assumptions made. **(8)**

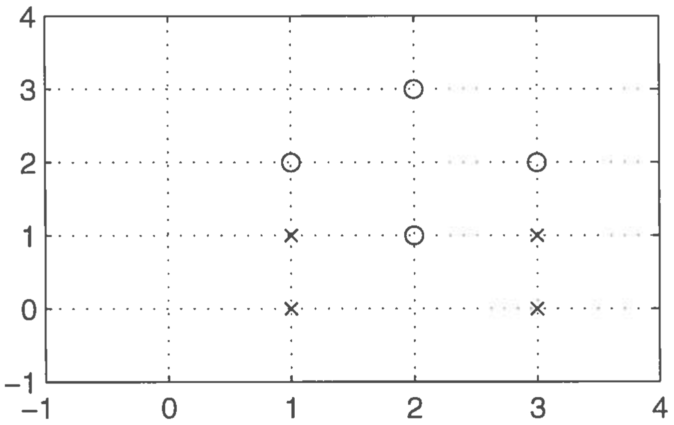

Figure B2a:

Question B3

a) Describe Principal Component Analysis (PCA), indicating its role in signal processing, the underlying optimization problem and its solution in terms of eigenvalue decomposition. **(5)**

b) Given the data: $x_1 = \binom{0}{0}, x_2 = \binom{2}{2}, x_3 = \binom{2}{4}, x_4 = \binom{4}{2}$ show that the Principal Directions are given by: $\binom{1/\sqrt{2}}{1/\sqrt{2}}$ and $\binom{1/\sqrt{2}}{-1/\sqrt{2}}$. Hence calculate the best 1-dimensional approximation of the data. **(8)**

c) Define the concept of Independent Component Analysis (ICA) indicating any key assumptions and ambiguities. **(3)**

d) A popular technique for ICA is to perform a nonlinear decorrelation of the data by diagonalizing the nonlinear correlation matrix, $\sum_n \varphi(\mathbf{s}_n)\mathbf{s}_n^T$, where $\mathbf{s}_n = \mathbf{W}\mathbf{x}_n$ are the independent sources and $\varphi(\mathbf{s}_n) = [\phi(s_n(1), \phi(s_n(2), \ldots, \phi(s_n(d)]$ is a threshold function applied element-wise with:

$$\phi(s_n(i)) = \begin{cases} 1 & \text{if } s_n(i) \geq 0 \\ -1 & \text{otherwise} \end{cases}$$

Determine whether the PCA decomposition given in part b) also provides a plausible ICA solution for the data in part b) when using this technique. **(4)**

END OF PAPER