UNIVERSITY COLLEGE LONDON

EXAMINATION FOR INTERNAL STUDENTS

MODULE CODE	:	**COMPGI13**
ASSESSMENT PATTERN	:	**COMPGI13B**
MODULE NAME	:	**Advanced Topics in Machine Learning**
DATE	:	**03-Jun-11**
TIME	:	**10:00**
TIME ALLOWED	:	**2 Hours 30 Minutes**

2010/11-COMPGI13B-001-EXAM-10

TURN OVER

Advanced Topics in Machine Learning, GI13, 2010/11

Answer ONE question from PART A and ONE question from PART B. Use separate answer books for each PART.

Marks for each part of each question are indicated in square brackets

Calculators are NOT permitted

Part A

1. a. Model the following problem with an MDP: you want to travel from London to Tokyo in the *shortest* time. Available connections are:
London → Tokyo; London → Frankfurt; Frankfurt → Tokyo; Frankfurt → Singapore; Singapore → Tokyo. Furthermore, the Singapore airport a probability p of being closed.

   - Define a Markov Decision Process.

     [5 marks]

   - Draw a directed graph which shows the states and the connections.

     [4 marks]

   - Define actions and write down P^a_{ij}, for state i, j and action a. (*Ignore zeros*).

     [3 marks]

   - A connection between state i and j has a travel time of t_{ij}. Use the travel time to define rewards and the discount (*the optimal control strategy is the one which maximises the value*).

     [5 marks]

   - Travel times t_{ij} are stochastic objects. Think of a way to incorporate that in your model. Assume you are travelling monthly from London to Tokyo and you do neither know the distribution of t_{ij} nor the distribution of p. How could you minimise the expected time it takes using your past experience (write down an estimator and define needed quantities).

     [8 marks]

b. Value approximation and optimisation:

- Write down the Bellman Equation for a given policy π.

  [3 marks]

- Write down an approximation scheme for V^{π} that is based on the Bellman Equation (*including initialisation and iteration steps*).

  [6 marks]

- Will the approximation scheme converge to the true value V^{π}? Why?

  [3 marks]

- Assume that you have a contraction operator A with contraction factor q for a metric $d(x,y)$ and assume that A has as a fixed-point at the value x^*. Write down the definition of a contraction. Assume you apply A iteratively and you are at iteration n. Bound the distance to the fixed-point in the metric d, i.e. bound $d(x^*,x^n)$ with $cd(x^{n+1},x^n)$, where $c \in \mathbb{R}$ is a constant. (*Hint: triangular inequality for d*).

  [8 marks]

- Assume you are looking for a fixed-point x^* in $\mathbb{R}^2$ and you have an operator A with $Ax^* = x^*$. It can happen that A is a contraction in one metric d_1 but not in a metric d_2. Interpret that statement. Construct a simple example where this happens (*Hint: d_1 the Euclidean distance; $d_2(x,y) := \max_{i\in\{1,2\}} |x_i - y_i|$. Take as a starting point the origin and as the fixed point $(10,10)$. Draw the behaviour of an operator A which is a contraction in d_1 but not in d_2.*)

  [5 marks]

[Total: 50 marks]

2. a. Explain the multi-armed bandit problem with time horizon T and L arms. What is known to the player and what is hidden. Explain the exploration-exploitation tradeoff.

[8 marks]

b. How can a Bandit problem, where the arms are related, be represented with a Gaussian Process? In many applications you are looking for the best arm given that your environment is in a certain state. Information about the state of the environment at time t is encoded in a feature vector ϕ_t. How can you use this information?

[8 marks]

c. Describe the Upper Confidence Bound (UCB) strategy indicating how it enables a trade-off between exploration and exploitation.

[10 marks]

d. Consider the situation described in part 2.b). Describe a UCB strategy making use of the Bayesian model updates, i.e. use the posterior mean and covariance equations to define a UCB rule.

[10 marks]

e. Describe an alternative method for selecting arms based on the posterior distribution obtained with the Bayesian model updates.

[8 marks]

f. Develop a simple approach where you use a Bandit algorithm to find an approximate solution for a Markov Decision Process. Assume that the transition matrix P and the reward R are unknown. Discuss the differences between this approach and an approach where you estimate the transition matrix, the reward distribution and where you use a dynamic programming algorithm to find a solution. Which algorithm will presumably be better, and why?

[6 marks]

[Total: 50 marks]

Part B

3. This question pertains to kernel methods and regularisation. Let $L: \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a loss function. Consider the optimisation problem

$$\min_{w \in \mathbb{R}^d} \left\{ \sum_{i=1}^{m} L(w^\top x_i, y_i) + \gamma \|w\|^2 \right\}, \tag{1}$$

where $w = (w_1, \ldots, w_d)$, $\|w\| = \sqrt{\sum_{i=1}^{d} w_i^2}$, γ is a positive number, and $(x_1, y_1), \ldots, (x_m, y_m) \in \mathbb{R}^d \times \mathbb{R}$ are given datapoints.

a. Assuming that, for every $y \in \mathbb{R}$, the function $t \mapsto L(t,y)$ is differentiable, give a proof of the Representer Theorem. That is, prove that the solution $\hat{w}$ is a linear combination of the datapoints $x_1, \ldots, x_m$.

[14 marks]

b. Consider now the case in which L is the square loss: $L(t,y) = (t-y)^2$. Use the representer theorem to reformulate problem (1) as the optimisation problem

$$\min_{c \in \mathbb{R}^m} \sum_{i=1}^{m} \left(y_i - \sum_{j=1}^{m} c_j x_j^\top x_i \right)^2 + \gamma \sum_{i,j=1}^{m} c_i c_j x_i^\top x_j$$

where $c = (c_1, \ldots, c_m)$. Which kind of optimisation problem do you obtain? Does it have a unique solution? explain your answer.

[12 marks]

c. Consider the kernel $K(x,t) := (1 + x^\top t)^2$ for $x, t \in \mathbb{R}^2$. Find a feature map associated to this kernel.

[12 marks]

d. Let A be a symmetric matrix. Show that the kernel $K(x,t) := x^\top A t$ is a valid kernel if and only if the matrix A is positive semidefinite.

[12 marks]

[total 50 marks]

4. This question pertains sparsity methods involving the ℓ_1 norm.

a. Consider the minimal ℓ_1-norm interpolation problem

$$\min\{\|w\|_1 : w \in \mathbb{R}^d,\ w^\top x_i = y_i,\ i = 1,\ldots,m\}$$

where $w = (w_1,\ldots,w_d)$ is a d-dimensional real vector, $\|w\|_1 := \sum_{i=1}^d |w_i|$, and $(x_1,y_1),\ \ldots,\ (x_m,y_m) \in \mathbb{R}^d \times \mathbb{R}$ are given datapoints.

Show that this problem is a linear programming problem.

[12 marks]

b. Consider the problem in the previous question when $m = 1$ and $y_1 > 0$. Give an example in which there is a unique solution and an example in which the solution is not unique.

[12 marks]

c. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ defined as

$$f(w_1,w_2) = \frac{1}{2}\left[(w_1-3)^2 + (w_2)^2\right] + \lambda(|w_1| + |w_2|) \qquad \text{for every } w_1, w_2 \in \mathbb{R},$$

where λ is a fixed positive constant. Argue that the above function is convex and the optimisation problem

$$\min\{f(w_1,w_2) : w_1, w_2 \in \mathbb{R}\} \tag{2}$$

is a quadratic programming problem.

[12 marks]

d. Argue that the solution of the above optimisation problem is unique; derive an explicit formula for the solution.

[14 marks]

[total 50 marks]

END OF PAPER