THE UNIVERSITY of ADELAIDE

Examination in School of Mathematical Sciences

Practice Exam Paper, 2018

105929 APP MTH 3020 Stochastic Decision Theory III

Official Reading Time:	10 mins
Writing Time:	120 mins
Total Duration:	130 mins

NUMBER OF QUESTIONS: 7 TOTAL MARKS: 73

Instructions

- Attempt all questions.
- Begin each answer on a new page.
- Examination materials must not be removed from the examination room.

Materials

- 1 Blue book is provided.
- Calculators without remote communications facilities are permitted.

DO NOT COMMENCE WRITING UNTIL INSTRUCTED TO DO SO.

1. Answer *true*, *false*, or *not necessarily true* to each of the following assertions. You must also provide a *very brief* (1–3 lines) justification for each of your answers. (You might, for example, wish to refer to a theorem discussed in lectures.)

   (a) For $x \in (-\infty, 0) \cup (0, \infty)$, $f(x) = x$ is a convex function.

   (b) Consider a mathematical program of the form:

   $$\begin{aligned} \text{minimise} \quad & z = g_0(\boldsymbol{x}) \\ \text{such that} \quad & g_i(\boldsymbol{x}) \leq 0,\ i = 1, \ldots, m \\ & \boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^n \quad . \end{aligned}$$

   A necessary and sufficient condition for the problem to be linear is the set $\mathcal{X}$ being a convex polytope.

   (c) In a recourse deterministic equivalent problem (DEP), if a solution $\boldsymbol{x}$ satisfies the first-stage constraints, then it will also satisfy the second-stage constraints.

   (d) A probabilistically constrained program is a generalisation of a recourse DEP.

   (e) Suppose we want to find the minimum walking distance from the Wayville Exam Hall to Ingkarni Wardli (IW). Suppose one optimal route is via the Original Cooper Alehouse (OCA). Then the part of that route from OCA to IW is also the shortest route for the sub-problem of minimising the walking distance from OCA to IW.

   (f) If a problem is either a discounted program, or a positive program, or a negative program, then we always have an optimal policy.

   (g) If we have a discounted program, then the optimal cost $F_s(i)$ of an s-horizon version converges to the optimal cost $F(i)$ of the infinite-horizon version, as $s \to \infty$.

   (h) The optimal cost for an infinite-horizon Markov decision problem is always defined.

   (i) We can always apply the Policy Iteration method to obtain the optimal policy in an infinite-horizon average-value MDP.

   (j) A hidden Markov model is a Markov chain in which the states are hidden.

[10 marks]

2. Maddy is having a party. She needs to buy enough pizza and champagne for everyone, assuming that each party attendee will eat one pizza and drink half a bottle of champagne. The number of people who will be present at her party, including Maddy herself, is a random variable d, where

$$d = \begin{cases} 20 & \text{w.p. } 0.2, \\ 17 & \text{w.p. } 0.7, \\ 1 & \text{w.p. } 0.1. \end{cases}$$

Each pizza costs \$10, and each bottle of champagne costs \$22. Each pizza or each bottle has to be bought as a whole. Maddy wants to minimise the cost of catering.

(a) Write down the stochastic linear program (SLP) for Maddy's catering problem.

(b) Write the naïve deterministic equivalent problem of the above SLP, and solve it.

(c) Suppose in the eventuality that Maddy has not bought in advance enough pizzas and champagne, she can quickly order more from the restaurant next door. In that case, the cost of a pizza is \$15, a bottle of champagne is \$30, and Maddy will no longer have to buy either pizza or champagne as a whole. Write the two-stage recourse DEP for Maddy's catering problem.

(d) Give the definition for a recourse matrix W of dimensions $m_1 \times n$ to be complete. In that case, is the recourse matrix in Part (c) complete? Justify.

(e) Give the definition for a recourse matrix W of dimensions $m_1 \times n$ to be relatively complete. In that case, is the recourse matrix in Part (c) relatively complete? Justify.

[17 marks]

3. (a) Give a necessary and sufficient condition for an $m_1 \times n$ recourse matrix W to be complete.

   (b) Consider the following recourse matrix

$$W = \begin{pmatrix} 1 & -1 & 1 & -1 & -5 \\ 0 & 1 & -1 & 0 & 10 \end{pmatrix}.$$

   Is W a complete recourse matrix? Justify.

   (c) The Dual Decomposition method, which solves DEPs with dual decomposition structure, involves the introduction of two types of constraints, *feasibility cuts* and *optimality cuts.* Explain the difference between the two types of constraints.

[8 marks]

4. Kate is contracted to run a restaurant called Orana over two years. At the end of each year, Orana is in one of the two states: $\{\text{POPULAR}, \text{AVERAGE}\} = \{1, 2\} =: \mathcal{S}$. At the beginning of each year, Kate makes a decision whether to renovate the restaurant.

   If Orana is POPULAR and Kate decides not to renovate, it stays POPULAR with probability (w.p.) 0.9 and becomes AVERAGE w.p. 0.1. If it is AVERAGE and she decides for no renovation, Orana stays AVERAGE w.p. 0.7 and becomes POPULAR w.p. 0.3.

   If Orana is POPULAR and Kate decides to renovate, it stays POPULAR w.p. 1. If it is AVERAGE and she decides to renovate, Orana stays AVERAGE w.p. 0.2 and becomes POPULAR w.p. 0.8.

   The cost of each renovation is \$0.5 million.

   If Orana remains POPULAR over the course of a year, its annual revenue is \$6 million. If it declines from POPULAR to AVERAGE, its annual revenue is \$4 million. On the other hand, if Orana remains AVERAGE, its annual revenue is 3 million dollars, and if it improves from AVERAGE to POPULAR, then its annual revenue is \$3.5 million.

   Assume that there is no advantage for Kate if Orana finishes at the end of the second year in any given state.

   Kate's objective is to maximise the expected revenue of the restaurant over the two years.

   (a) Write the transition probability matrices for Orana under each of Kate's decisions.

   (b) Write the reward matrices for Orana under each of Kate's decisions.

   (c) Let $V_k(i)$ denote the optimal expected revenue, given Orana is in state $i \in \mathcal{S}$ at the beginning of stage k, for $k = 0, 1, 2$. Write the dynamic programming equation to obtain $V_k(i)$, and relevant boundary conditions.

   (d) Compute $V_1(\text{POPULAR})$, and state the optimal action. Justify.

[10 marks]

5. A collection of n songs is to be played in arbitrary order by a disk jockey at a local radio station. Song i has playing time p_i, and when it completes a reward r_i is obtained from the sponsor of the song. The rewards are discounted by parameter $\beta \in (0,1)$ as the earlier in the program the songs are completed, the more valuable they are for their respective sponsor. We assume that the songs are played in the order $i_1, i_2, \ldots, i_k, i, j, i_{k+3}, \ldots, i_{n-1}, i_n$.

   (a) Noting that the decision points are not equally spaced in time, explain why the reward under this schedule is given by

   $$\beta^{p_{i_1}} r_{i_1} + \beta^{p_{i_1}+p_{i_2}} r_{i_2} + \cdots + \beta^{T+p_i} r_i + \beta^{T+p_i+p_j} r_j + \cdots + \beta^N r_n,$$

   where $T = p_{i_1} + p_{i_2} + \cdots + p_{i_k}$ and $N = \sum_{j=1}^n p_j$.

   (b) If the play order of songs i and j are reversed, give an expression in terms of the ratio $\dfrac{r_i \beta^{p_i}}{1-\beta^{p_i}}$ that enables you to choose one play order over the other.

   (c) Using an interchange argument, find the order of playing that maximises the sum of discounted rewards.

[10 marks]

6. Given a model $(\boldsymbol{\theta}_p, \boldsymbol{\phi}_p, \boldsymbol{p}_0)$, and a sequence of observations $\boldsymbol{y}$ of the random variable Y, where

   - $\boldsymbol{\theta}_p$ are the parameters of a homogeneous Markov chain $\{X_t \big| t \geq 0\}$ that describes the state transitions of a Hidden Markov chain $P(X_t = j \big| X_{t-1} = i)$.
   - $\boldsymbol{\phi}_p$ are the parameters of the observation probability mass function $P(Y_t = y \big| X_t = i)$.
   - $\boldsymbol{p}_0$ is the initial state distribution $P(X_0 = i)$.

   Let $\Lambda = (\boldsymbol{\theta}_p, \boldsymbol{\phi}_p, \boldsymbol{p}_0)$, $\boldsymbol{y}_T = (y_0, y_1, \ldots, y_T)$ be the observation sequence, and $\boldsymbol{x}_T = (x_0, x_1, \ldots, x_T)$ be the state sequence.

   (a) Give the joint probability of $\boldsymbol{x}_T$ and $\boldsymbol{y}_T$, given Λ. Justify.

   (b) Using the above result to obtain the probability of observing $\boldsymbol{y}_T$ given Λ.

   (c) Explain why you would not in general use the above expression to find $\Pr(\boldsymbol{y} \mid \Lambda)$.

   (d) Which algorithm would you use instead? Give a brief description of that algorithm.

   (e) Prove that the algorithm gives the correct answer. *Hint:* Prove that Step 2 works!

   [18 marks]