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.

   False. Because while for all $x, y \in \mathbb{R}$ and for $\lambda \in (0, 1)$, we have

   $$\lambda x + (1 - \lambda)y = \lambda x + (1 - \lambda)y,$$

   the set $(-\infty, 0) \cup (0, \infty)$ on which $f(x)$ is defined is not convex. [1 mark]

   (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.

   False. The problem is said to be linear if $\mathcal{X}$ is a convex polytope and the functions g_i are linear, for $i = 1, \ldots, n$. [1 mark]

   (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.

   False. This is only true if the recourse DEP has the relatively complete recourse property. [1 mark]

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

   False. A probabilistically constrained program requires that the stochastic constraints are satisfied for at least some given probability $\alpha > 0$; a recourse deterministic equivalent problem allows the stochastic constraints to be violated (for example, with probability 1!). [1 mark]

   (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.

   True. This is an example of the Principle of Optimality. [1 mark]

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

   False. An example is the positive program considered in class, on the set of integers, where there are two actions, of going to state $i + 1$ with no reward or state 0 with reward $1 - 1/i$. [1 mark]

Please turn over for page 3

(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 $s \to \infty$.

True. This was proven in class. (The same applies for positive programs, only negative programs require an additional condition.) [1 mark]

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

False. For example, consider a simple case of the cost at each step being 1 in the odd steps, and -1 in the even ones. [1 mark]

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

False. This works only if there exists an average-value per stage g, given by

$$g = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{n=0}^{T} v(X_n, f_n(X_n))\right].$$

[1 mark]

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

False. A hidden Markov model is a doubly stochastic process, in which the underlying process is hidden and the other process, which depends on the underlying one, is observed. [1 mark]

[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.

Let x_1 be the number of pizzas and x_2 be the number of bottles of champagne Maddy needs to buy. Then, the stochastic LP for Maddy's catering problem is

$$\begin{aligned} \min \quad & 10x_1 + 22x_2 \\ \text{subject to} \quad & x_1 \geq d \\ & 2x_2 \geq d \\ & x_1, x_2 \geq 0 \\ & x_1, x_2 \text{ are integers.} \end{aligned}$$

[4 marks]

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

$$\begin{aligned} \min \quad & 10x_1 + 22x_2 \\ \text{subject to} \quad & x_1 \geq 20 \\ & x_1 \geq 17 \\ & x_1 \geq 1 \\ & 2x_2 \geq 20 \\ & 2x_2 \geq 17 \\ & 2x_2 \geq 1 \\ & x_1, x_2 \geq 0 \\ & x_1, x_2 \text{ are integers.} \end{aligned}$$

Clearly, four of the above constraints are redundant:

$$x_1 \geq 17, \quad x_1 \geq 1, \quad 2x_2 \geq 17, \quad 2x_2 \geq 1.$$

By inspection, it's clear that $x_1 = 20$ and $x_2 = 10$, with the optimal cost being $\$200 + \$220 = \$420$. [4 marks]

(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.

$$\begin{aligned} \min \quad & 10x_1 + 22x_2 + \mathbb{E}_d[Q(x,d)] \\ \text{subject to} \quad & x_1 \geq d \\ & 2x_2 \geq d \\ & x_1, x_2 \geq 0, \\ & x_1, x_2 \text{ are integers,} \end{aligned}$$

where

$$\mathbb{E}_d[Q(x,20)] = 0.2Q(x,20) + 0.9Q(x,17) + 0.1Q(x,1),$$

and

$$Q(x,20) = \min_{y_1,y_2} \left\{ 15y_1 + 30y_2 : \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 20 - x_1 \\ 20 - 2x_2 \end{bmatrix}, \quad y_1, y_2 \geq 0 \right\}$$
$$Q(x,17) = \min_{y_1,y_2} \left\{ 15y_1 + 30y_2 : \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 17 - x_1 \\ 17 - 2x_2 \end{bmatrix}, \quad y_1, y_2 \geq 0 \right\}$$
$$Q(x,1) = \min_{y_1,y_2} \left\{ 15y_1 + 30y_2 : \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 - x_1 \\ 1 - 2x_2 \end{bmatrix}, \quad y_1, y_2 \geq 0 \right\}$$

[5 marks]

(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.

A recourse matrix W of dimensions $m_1 \times n$ is said to be complete if

$$\{\boldsymbol{z} : \boldsymbol{z} = W\boldsymbol{y}, \quad \boldsymbol{y} \geq \boldsymbol{0}\} = \mathbb{R}^{m_1}.$$

In the above recourse problem, we have $W = I$, which is clearly not complete. For example, pick $\boldsymbol{z} = (-1, 0)$. [2 marks]

(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.

A recourse matrix W of dimensions $m_1 \times n$ is said to be relatively complete if for every feasible first-stage solution $\boldsymbol{x}$ the second-stage program is feasible.

In the above recourse problem, for every feasible first-stage solution, we can pick $\boldsymbol{y}$ to be precisely the absolute values of the constraint violations. [2 marks]

[17 marks]

Please turn over for page 6

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

An $m_1 \times n$ matrix W is a complete recourse matrix if and only if

1. it has $\text{rank}(W) = m_1$, and
2. assuming without loss of generality that its first m_1 columns $W_{:,1}, \dots, W_{:,m_1}$ are linearly independent, the linear constraints

$$\begin{aligned} W\boldsymbol{y} &= \mathbf{0} \\ y_i &\geq 1 \quad \text{for } i = 1, \dots, m_1 \\ \boldsymbol{y} &\geq \mathbf{0} \end{aligned}$$

have a feasible solution.

[3 marks]

(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.

As the first two columns of W are linearly independent, $\text{rank}(W) = 2$.

Then, for

$$\boldsymbol{y} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \end{bmatrix}^\top \geq \mathbf{0},$$

we see that $y_1, y_2 \geq 1$ and $y_3, y_4, y_5 \geq 0$. This, together with the fact that $W\boldsymbol{y} = \mathbf{0}$, implies that W is a complete recourse matrix. [3 marks]

(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.

A feasibility cut is a linear constraint introduced to cut off infeasible solutions to the second-stage program; an optimality cut is a linear constraint introduced to cut off solutions that will never be optimal for the DEP. [2 marks]

[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.

$$P(\text{NO}) = \begin{bmatrix} 0.9 & 0.1 \\ 0.3 & 0.7 \end{bmatrix}, \quad P(\text{RENOVATE}) = \begin{bmatrix} 1 & 0 \\ 0.8 & 0.2 \end{bmatrix}.$$

[2 marks]

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

$$R(\text{NO}) = \begin{bmatrix} 6 & 4 \\ 3.5 & 3 \end{bmatrix}, \quad R(\text{RENOVATE}) = R(\text{NO}) - 0.5 = \begin{bmatrix} 5.5 & 3.5 \\ 3 & 2.5 \end{bmatrix}.$$

[2 marks]

   (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.

   The boundary conditions are

$$V_2(\text{POPULAR}) = V_2(\text{AVERAGE}) = 0.$$

   The dynamic programming equation is

$$V_k(i) = \max_{u_k \in \{\text{NO}, \text{RENOVATE}\}} \left\{ c(i, u_k) + \sum_{j=1,2} p_{ij}(u_k) V_{k+1}(j) \right\},$$

   where $c(i, u_k)$ is the expected reward for taking action u_k in i, for $k = 0, 1$. [3 marks]

(d) Compute $V_1(\textsc{popular})$, and state the optimal action. Justify.
First,

$$\begin{aligned} V_1(\textsc{popular}) &= \min\{\text{optimal ER under } \textsc{no}, \quad \text{optimal ER under } \textsc{renovate}\} \\ &= \min\{0.9 \times 6 + 0.1 \times 4 + 0, \quad 5.5 + 0\} \\ &= \min\{5.8, 5.5\}. \end{aligned}$$

So the optimal action is NO, with the optimal value \$5.8 million. [3 marks]

[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 $N = \sum_{j=1}^{n} p_j$.

   The discounting applies to the reward from the start until the time the song is played and hence we discount the reward for the first song played by the time it takes to play. Similarly, the second song finishes playing at time $p_{i_1} + p_{i_2}$ and is discounted by $\beta^{p_{i_1}+p_{i_2}}$, etc. [4 marks]

   (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.

   If we let $R_1 = \beta^{p_{i_1}} r_{i_1} + \beta^{p_{i_1}+p_{i_2}} r_{i_2} + \cdots + \beta^T r_{i_k}$ and $R_2 = \beta^{T+p_i+p_j+p_{i_{k+3}}} r_{i_{k+3}} + \cdots + \beta^N r_n$, then the rewards under the two schedules are respectively

   $$R_1 + \beta^{T+p_i} r_i + \beta^{T+p_i+p_j} r_j + R_2 \quad \text{and} \quad R_1 + \beta^{T+p_j} r_j + \beta^{T+p_j+p_i} r_i + R_2,$$

   so that we choose to play song i before song j if

   $$\frac{r_i \beta^{p_i}}{1-\beta^{p_i}} > \frac{r_j \beta^{p_j}}{1-\beta^{p_j}}.$$

   [4 marks]

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

   Since this pair of songs is arbitrary, we can order them in decreasing order of the indices $\dfrac{r_i \beta^{p_i}}{1-\beta^{p_i}}$ and play them in that order. [2 marks]

[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.

   $$\begin{aligned}&\Pr\left(\boldsymbol{X}_T = \boldsymbol{x}_T, \boldsymbol{Y}_T = \boldsymbol{y}_T \mid \Lambda\right)\\ &= \Pr\left(\boldsymbol{X}_T = \boldsymbol{x}_T \mid \Lambda\right) \Pr\left(\boldsymbol{Y}_T = \boldsymbol{y}_T | \boldsymbol{X}_T = \boldsymbol{x}_T, \Lambda\right)\\ &= \underbrace{\Pr(X_0 = x_0) \prod_{i=1}^{t} \Pr(X_i = x_i | X_{i-1} = x_{i-1})} \overbrace{\prod_{j=0}^{t} \Pr(Y_j = y_j | X_j = x_j)},\end{aligned}$$

   where the first product (underbraced) is based on the Markov property, and the second product (overbraced) is based on the conditional independence property of the observation random variables $\{Y_n\}$. [5 marks]

   (b) Using the above result to obtain the probability of observing $\boldsymbol{y}_T$ given Λ.
   Summing over all state sequences $\boldsymbol{x}$, we get

   $$\begin{aligned}\Pr(\boldsymbol{y} \big| \Lambda) &= \sum_{\boldsymbol{x}} p(\boldsymbol{y}, \boldsymbol{x} \big| \Lambda)\\ &= \sum_{x_0=1}^{N} \sum_{x_1=1}^{N} \cdots \sum_{x_T=1}^{N} p_{x_0} \prod_{i=1}^{T} p_{x_{i-1}, x_i} \prod_{j=0}^{T} P(y_j | x_j).\end{aligned}$$

   [1 mark]

   (c) Explain why you would not in general use the above expression to find $\Pr(\boldsymbol{y} \mid \Lambda)$.
   The naive approach is generally infeasible, since it requires $(2T + 1)N^{T+1}$ multiplications. [1 mark]

   (d) Which algorithm would you use instead? Give a brief description of that algorithm.
   To find $\Pr(\boldsymbol{y} \mid \Lambda)$, we use the so called "$\alpha$-pass" or forward algorithm, defined by the forward function

   $$\alpha_t(i) = P(y_0, y_1, \ldots, y_t, X_t = i \big| \Lambda), \text{ for } t = 0, 1, \ldots, T \text{ and } i = 1, 2, \ldots, N,$$

   where $\alpha_t(i)$ is the probability of the partial observation sequence up to time t, where the underlying Markov process is in state i at time t, which can be computed recursively as follows

1. For $i = 1, 2, \ldots, N$, let

$$\alpha_0(i) = P(y_0, X_0 = i \big| \Lambda) = P(X_0 = i) P(y_0 \big| X_0 = i, \Lambda).$$

2. For $t = 1, 2, \ldots, T$ and $i = 1, 2, \ldots, N$ we then compute

$$\alpha_t(i) = \left(\sum_{j=1}^{N} \alpha_{t-1}(j)\, p_{ji} \right) P(y_t \big| X_t = i).$$

3. Then from part (b) we see that

$$P(\boldsymbol{y} \big| \Lambda) \;=\; \sum_{i=1}^{N} \alpha_T(i).$$

[6 marks]

(e) Prove that the algorithm gives the correct answer. *Hint:* Prove that Step 2 works!
The only part we need to prove

$$\begin{aligned}
a_t(i) &= \Pr\left(\boldsymbol{Y}_t = \boldsymbol{y}_t, X_t = i\right) \\
&= \sum_{j=1}^{N} \Pr\left(\boldsymbol{Y}_t = \boldsymbol{y}_t, X_t = i, X_{t-1} = j\right) \\
&= \sum_{j=1}^{N} \Pr\left(Y_t = y_t, X_t = i \mid \boldsymbol{Y}_{t-1} = \boldsymbol{y}_{t-1}, X_{t-1} = j\right) \Pr\left(\boldsymbol{Y}_{t-1} = \boldsymbol{y}_{t-1}, X_{t-1} = j\right) \\
&= \sum_{j=1}^{N} \Pr\left(y_t \mid \boldsymbol{y}_{t-1}, X_{t-1} = j, X_t = i\right) \Pr\left(X_t = i \mid \boldsymbol{y}_{t-1}, X_{t-1} = j\right) \Pr\left(\boldsymbol{y}_{t-1}, X_{t-1} = j\right) \\
&=^* \Pr\left(y_t \mid X_t = i\right) \sum_{j=1}^{N} \Pr\left(X_t = i \mid X_{t-1} = j\right) \Pr\left(\boldsymbol{y}_{t-1}, X_{t-1} = i\right) \\
&= \left(\sum_{j=1}^{N} \alpha_{t-1}(j)\, p_{ji} \right) \Pr(y_t \mid X_t = i).
\end{aligned}$$

Note that for $=^*$, we have used the conditional independence property of the hidden Markov chain, and the Markov property of $\{X_t\}$. [5 marks]

[18 marks]