

**Examination in School of Mathematical Sciences
Semester 2, 2015**

105929 APP MTH 3020 Stochastic Decision Theory III

Official Reading Time: 10 mins

Writing Time: 120 mins

Total Duration: 130 mins

NUMBER OF QUESTIONS: 6 TOTAL MARKS: 70

Instructions

- Attempt all questions.
- This examination comprises 70% of the total assessment for this course.
- Begin each answer on a new page.
- Examination materials must not be removed from the examination room.

Materials

- 1 Blue book is provided.
- Formulae sheets are provided.
- Calculators without remote communications or CAS capability are allowed.

DO NOT COMMENCE WRITING UNTIL INSTRUCTED TO DO SO.

Stochastic Linear Programming

1. A white goods manufacturer makes freezers and fridges, where due to the physical layout of the premises, there is a total production limit of 1500 units of white goods per production period. Both assembly and finishing are required resources for the manufacturing process. The requirements measured in hours per unit are known and shown in the table below along with the profit per unit of product.

Product parameters	freezers	fridges
Assembly hours	8	16
Finishing hours	6	4
profit per unit	\$150	\$250

Our problem is to select the product mix to maximise total profit when the availability of the labour resources are not known. Rather we have two equally likely realisations of the hours available for manufacture.

$$\epsilon_1 = \begin{cases} 4950 & \text{assembly hours} \\ 3636 & \text{finishing hours} \end{cases}$$

$$\epsilon_2 = \begin{cases} 5850 & \text{assembly hours} \\ 4064 & \text{finishing hours} \end{cases}$$

For the example, the expected time available for assembly is 5400 hours and the expected time available for finishing is 3850 hours.

- (a) Write down a LP in terms of the two primal variables x_1, x_2 , where x_1 is the number of freezer units and x_2 is the number of fridge units produced. This LP should be able to be used to find the maximum profit when the expected available times for assembly and finishing are used to define the problem constraints.

$$\begin{aligned} \text{(P) max } & z = 150x_1 + 250x_2 \\ \text{s.t. } & 8x_1 + 16x_2 \leq 5400 \\ & 6x_1 + 4x_2 \leq 3850 \\ & x_1 + x_2 \leq 1500 \\ & x_1, x_2 \geq 0 \end{aligned}$$

[4 marks]

Please turn over for page 3

- (b) Verify that 625 freezers and 25 fridges is a feasible solution of the original LP and find the profit.

The \mathbf{x} satisfies all constraints and $z = 150 \times 625 + 250 \times 25 = \$100,000$.

[2 marks]

- (c) Write down the dual LP in terms of dual variables y_1, y_2, y_3 .

$$\begin{aligned} \text{(D) min } z &= 5400y_1 + 3850y_2 + 1500y_3 \\ \text{s.t. } 8y_1 + 6y_2 + y_3 &\geq 150 \\ 16y_1 + 4y_2 + y_3 &\geq 250 \\ y_1, y_2, y_3 &\geq 0 \end{aligned}$$

[4 marks]

- (d) Verify that $y_1 = 14.0625, y_2 = 6.25, y_3 = 0$ is a feasible solution to the dual LP and find the value of the objective function.

The \mathbf{y} satisfies all constraints and $5400 \times 14.0625 + 3850 \times 6.25 + 1500 \times 0 = \$100,000$.

[2 marks]

- (e) Giving an explanation, are these solutions optimal in each case?

As the objective functions for both the primal and dual are the same, then by the Theorem of Strong Duality, the solution is optimal for both LPs.

[2 marks]

[14 marks]

Stochastic Linear Programming

2. A solution to the averaged value LP in the previous question is not very acceptable because it does not allow for the stochastic variation of available assembly and finishing hours.

Assume that additional assembly hours may be purchased at \$50 per hour and that extra finishing hours may be purchased at \$120 per hour.

Also assume that any unused base assembly hours are wasted and must be costed at \$40 per hour and similarly any unused base hours of finishing must be costed at \$90 per hour.

- (a) Giving some explanation, write down the expanded version of a recourse model considering both realisations as outlined.

Hint: You will have a \mathbf{y} vector of length $2 \times 4 = 8$.

$$\begin{aligned}
 \max \quad & z = 150x_1 + 250x_2 - \frac{1}{2}(40y_1 + 50y_2 + 40y_3 + 50y_4 + 90y_5 + 120y_6 + 90y_7 + 120y_8) \\
 \text{s.t.} \quad & 8x_1 + 16x_2 + y_1 - y_2 = 4950 \\
 & 8x_1 + 16x_2 + y_3 - y_4 = 5850 \\
 & 6x_1 + 4x_2 + y_5 - y_6 = 3636 \\
 & 6x_1 + 4x_2 + y_7 - y_8 = 4064 \\
 & x_1 + x_2 \leq 1500 \\
 & x_1, x_2 \geq 0, y_i \geq 0.
 \end{aligned}$$

Note that the y_i variables for i odd, correspond to wasted hours as they take up the slack and the y_j variables for j even, correspond to purchasing extra hours.

[4 marks]

(b) Define $A, \mathbf{b}, \mathbf{q}^T, W, T(\boldsymbol{\xi})$ and $\mathbf{h}(\boldsymbol{\xi})$, to rewrite the above problem in the form

$$\begin{aligned}
 \max \quad & z = \mathbf{c}^T \mathbf{x} - E[Q(\mathbf{x}, \boldsymbol{\xi})]_{\boldsymbol{\varepsilon}_i} \\
 \text{s.t.} \quad & A\mathbf{x} \leq \mathbf{b} \\
 & \mathbf{x} \geq \mathbf{0}.
 \end{aligned}$$

with $Q(\mathbf{x}, \boldsymbol{\varepsilon}_i)$ for each realisation $\boldsymbol{\varepsilon}_i$ of $\boldsymbol{\xi}$ given by

$$\begin{aligned}
 \min \quad & \mathbf{q}^T \mathbf{y} \\
 \text{s.t.} \quad & W\mathbf{y} = \mathbf{h}(\boldsymbol{\varepsilon}_i) - T(\boldsymbol{\varepsilon}_i)\mathbf{x}, \\
 & \mathbf{y} \geq \mathbf{0},
 \end{aligned}$$

where $\mathbf{y}^T = (y_1, y_2, y_3, y_4)$.

Note that $A = \begin{bmatrix} 1 & 1 \end{bmatrix}$, $\mathbf{b} = 1500$, so that equivalently we have

$$\begin{aligned}
 \max \quad & z = 150x_1 + 250x_2 - E[Q(\mathbf{x}, \boldsymbol{\xi})]_{\boldsymbol{\xi}} \\
 \text{s.t.} \quad & x_1 + x_2 \leq 1500 \\
 & x_1, x_2 \geq 0,
 \end{aligned}$$

where $\mathbf{x} = (x_1, x_2)^T$, $\mathbf{y} = (y_1, y_2, y_3, y_4)^T$,

$\mathbf{c}^T = (150, 250)$, $\mathbf{q}^T = (40, 50, 90, 120)$,

$$W = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \quad T(\boldsymbol{\varepsilon}_1) = T(\boldsymbol{\varepsilon}_2) = \begin{bmatrix} 8 & 16 \\ 6 & 4 \end{bmatrix},$$

$\mathbf{h}(\boldsymbol{\varepsilon}_1) = (4950, 3636)^T$ and $\mathbf{h}(\boldsymbol{\varepsilon}_2) = (5850, 4064)^T$.

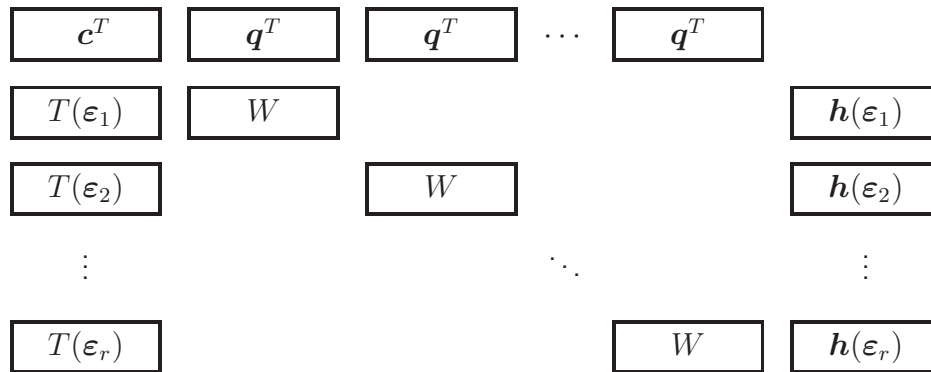
[4 marks]

- (c) Giving an explanation, is W a complete recourse matrix?

W has independent rows, which implies that it has rank 2 and hence $W\mathbf{y} = \mathbf{h}(\boldsymbol{\varepsilon}_i) - T(\boldsymbol{\varepsilon}_i)\mathbf{x}$ has non-trivial solutions $\mathbf{y} \geq \mathbf{0}$ for all first stage solutions \mathbf{x} . Hence we say that W is of complete recourse. [2 marks]

- (d) Give a brief explanation of how the L-shaped algorithm is beneficial in finding a solution to these types of problem.

The computational burden associated with the number r of realisations $\boldsymbol{\varepsilon}$ of the random vector $\boldsymbol{\xi}$ can make the above expanded form intractable, but this can be significantly reduced by associating one set of second-stage decisions vectors \mathbf{y}_k to each realisation $\mathbf{q}_k, \mathbf{h}_k, T_k$ for $1 \leq k \leq r$ in an iterative process. The idea can be viewed in the following diagram, where for each realisation $\boldsymbol{\varepsilon}_k$ of $\boldsymbol{\xi}$, for $1 \leq k \leq r$, there is an extra row. The L-shaped algorithm takes advantage of this structure, solving for the



first stage variables \mathbf{x} and then iteratively solving second stage LPs associated with each row. During this iterative process the L-shaped algorithm considers first the feasibility of the problem and adds feasibility cuts to restrict the first stage solutions so that the second stage will have feasible solutions. After this is guaranteed (assuming a solution exists), the algorithm adds optimality constraints (if required) to the first stage problem to eventually achieve the optimal solution. These LPs are significantly smaller than the expanded form and are easily solved and yet this process effects the same solution to the larger problem. [4 marks]

The initial first stage problem has the form

$$\begin{aligned}
 \max \quad z &= 150x_1 + 250x_2 \\
 \text{s.t.} \quad &x_1 + x_2 \leq 1500 \\
 &x_1, x_2 \geq 0,
 \end{aligned}$$

which trivially yields $\mathbf{x} = (0, 1500)$. Then we have at the first step of the Optimality

cuts part of the algorithm the following LP to solve

$$\begin{array}{ll}\min & 40y_1 + 50y_2 + 90y_3 + 120y_4 \\ \text{s.t.} & y_1 - y_2 = 4950 - 8(0) - 16(1500) = -19050 \\ & y_3 - y_4 = 3636 - 6(0) - 4(1500) = -2364 \\ & \mathbf{y} \geq \mathbf{0},\end{array}$$

[6 marks]

[14 marks]

Markov Decision Problems

3. Consider a finite state Markov Decision Process, where for some chosen policy, over a finite horizon T we have a value function given by

$$V_k(i) = \sum_{j=1}^n p_{i,j} (r_{i,j} + V_{k+1}(j)), \quad (1)$$

on states $i \in \{1, \dots, n\}$, where $r_{i,j}$ is a reward associated with a transition from state i to state j with probability $p_{i,j}$ under this chosen policy. Assume also that for $T \gg k$, we have that the process is such that there exists a positive constant g such that for each $i \in \{1, \dots, n\}$

$$V_k(i) = g \times (T - k) + v_i, \quad (2)$$

where v_i is the value of starting in state i at time k .

- (a) Using the above information show that for each $i \in \{1, \dots, n\}$

$$v_i + g = \left(\sum_{j=1}^n p_{i,j} r_{i,j} \right) + \left(\sum_{j=1}^n p_{i,j} v_j \right).$$

Substituting (2) into (1) yields

$$\begin{aligned} (T - k)g + v_i &= \sum_{j=1}^n p_{i,j} (r_{i,j} + (T - k - 1)g + v_j) \\ \implies (T - k)g + v_i - (T - k - 1)g &= \sum_{j=1}^n p_{i,j} (r_{i,j} + v_j) \end{aligned}$$

or

$$v_i + g = \left(\sum_{j=1}^n p_{i,j} r_{i,j} \right) + \left(\sum_{j=1}^n p_{i,j} v_j \right).$$

as required.

[4 marks]

- (b) Explain how the equations in part (a) may be used to define the policy improvement routine for each state i .

These equations represent the value determination operation. The policy improvement routine for each state i , is to find an alternative policy that maximises the expression on the RHS of this equation. They are a set of n linear equations in $n + 1$ unknowns v_1, v_2, \dots, v_n and g . A unique solution may be found by setting $v_1 = 0$, so that the other unknowns, which represent relative values of starting in $i \neq 1$, can be found.

The next step is to apply the PIR criteria by trying an alternative policy in the RHS. The PIR will generally not immediately yield the optimal policy unless our first guess (initial policy) was an exceptionally good choice, so that there is no alternative policy such that the RHS yields an improved value.

[4 marks]

[8 marks]

Markov Decision Problems

4. A co-op produces toys and at the end each year may be in one of two states $X \in \{1, 2\}$, where 1 is a healthy state and 2 is an unhealthy state. At the beginning of each year, they can take a decision to research the market overseas (at a cost of 4) or not such that the state of the co-op and a reward is given by the following state transition and reward matrices

$$\begin{aligned} P(\text{research}) &= \begin{pmatrix} 0.75 & 0.25 \\ 0.8 & 0.2 \end{pmatrix} \quad \text{and} \quad P(\text{no research}) = \begin{pmatrix} 0.5 & 0.5 \\ 0.2 & 0.8 \end{pmatrix} \\ R(\text{research}) &= \begin{pmatrix} 4 & -2 \\ -2 & -8 \end{pmatrix} \quad \text{and} \quad R(\text{no research}) = \begin{pmatrix} 8 & 2 \\ 2 & -4 \end{pmatrix}. \end{aligned}$$

- (a) Giving explanation, write down a Bellman equation for the expected monetary value $V_0^\Psi(i)$ for the co-op under a policy Ψ starting from state $i \in \{1, 2\}$ at the beginning of stage $k = 0$ over a finite horizon T .

$$V_0^\Psi(i) = E \left[V_T(X_T^\Psi) + \sum_{n=0}^{T-1} v(X_n^\Psi, \mu(X_n^\Psi)) \right]_\Psi,$$

where $X_0^\Psi = i$, $V_T(X_T^\Psi)$ is the value received for finishing in state X_T^Ψ and $v(X_n^\Psi, \mu(X_n^\Psi))$ is the single stage expected reward given that the co-op starts in state X_n , at the beginning of stage n under decision (action) $\mu(X_n^\Psi)$.

[4 marks]

- (b) Perform a value iteration procedure on this problem to find the optimal policy and value over a time horizon of $T = 4$ years for starting in state $i \in \{1, 2\}$, assuming that there is a zero reward for finishing in either state. That is, $V_4(1) = V_4(2) = 0$.

According to the Bellman optimisation principle, we maximise the above equation at each stage as we proceed backward in time from stage $T - 1$ to stage 0.

$$\begin{aligned} V_4^*(1) &= V_4^*(2) = 0 \\ V_3^*(1) &= \max\{0.5(8 + 0) + 0.5(2 + 0), 0.75(4 + 0) + 0.25(-2 + 0)\} = 5 \quad \text{no} \\ V_3^*(2) &= \max\{0.2(2 + 0) + 0.8(-4 + 0), 0.8(-2 + 0) + 0.2(-8 + 0)\} = -2.8 \quad \text{no} \\ V_2^*(1) &= \max\{0.5(8 + 5) + 0.5(2 - 2.8), \\ &\quad 0.75(4 + 5) + 0.25(-2 - 2.8)\} = 6.1 \quad \text{no} \\ V_2^*(2) &= \max\{0.2(2 + 5) + 0.8(-2 - 2.8), \\ &\quad 0.8(-2 + 5) + 0.2(-8 - 2.8)\} = 0.24 \quad \text{res} \\ V_1^*(1) &= 8.17 \quad \text{no} \\ V_1^*(2) &= 1.728 \quad \text{res} \\ V_0^*(1) &= 9.949 \quad \text{no} \\ V_0^*(2) &= 3.6816 \quad \text{res} \end{aligned}$$

[8 marks]

- (c) From the value iteration you have just performed what would you propose as an optimal stationary policy if the co-op was to run indefinitely?

The optimal stationary policy would be to always do research if in state 2 and never do research if in state 1.

[2 marks]

[14 marks]

5. The Greek adventurer Theseus is trapped in a room from which lead n passages. Theseus knows that if he enters passage i for $i \in \{1, \dots, n\}$, one of three fates will befall him:

- he will escape with probability p_i ,
- he will be killed with probability q_i , and
- with probability $r_i = 1 - p_i - q_i$ he will find the passage to be a dead end and be forced to return to the room.

The fates associated with different passages are independent. Using an interchange argument, establish the order in which Theseus should attempt the passages if he wishes to maximise his probability of eventual escape.

Consider that Theseus can try the passages in some order

$$d_1, d_2, \dots, d_i, d_j, d_k, d_\ell, \dots, d_n,$$

and then consider if instead, he adopted the order

$$d_1, d_2, \dots, d_i, d_k, d_j, d_\ell, \dots, d_n,$$

where the passages d_j and d_k have been swapped. The outcome up to and including passage i is identical. The respective probabilities of escape for the above scenarios from that point on are

$$p_{d_j} + r_{d_j}p_{d_k} + r_{d_j}r_{d_k}P_{\ell, \dots, n} \quad \text{and} \quad p_{d_k} + r_{d_k}p_{d_j} + r_{d_k}r_{d_j}P_{\ell, \dots, n},$$

where $P_{\ell, \dots, n}$ is the probability of escape when trying the remaining passages in order d_ℓ, \dots, d_n . Hence, we would choose to try passage j before passage k if $p_{d_j} + r_{d_j}p_{d_k} \geq p_{d_k} + r_{d_k}p_{d_j}$, reducing to,

$$\frac{p_{d_j}}{p_{d_j} + q_{d_j}} \geq \frac{p_{d_k}}{p_{d_k} + q_{d_k}}$$

or equivalently,

$$\frac{p_{d_j}}{q_{d_j}} \geq \frac{p_{d_k}}{q_{d_k}}.$$

As this is independent of where it occurred in the sequence, he should try passages in order of decreasing value of p_i/q_i .

[8 marks]

Hidden Markov Models

6. (a) Given a model (θ_p, ϕ_p, p_0) , and a sequence of observations \mathbf{O} of the random variable Y , where
- θ_p are the parameters of a homogeneous Markov chain $\{X_t | t \geq 0\}$ that describes the state transitions of a Hidden Markov chain $P(X_t = j | X_{t-1} = i)$.
 - ϕ_p are the parameters of the observation probability mass function $P(Y_t = y | X_t = i)$.
 - p_0 is the initial state distribution $P(X_0 = i)$.

Describe the three classical problems associated with the above hidden Markov model.

1. Given the model (θ_p, ϕ_p, p_0) and a sequence of observations \mathbf{O} , determine the likelihood of the observation sequence \mathbf{O} .
2. Given the model (θ_p, ϕ_p, p_0) and a sequence of observations \mathbf{O} , find an optimal state sequence for the underlying (hidden) Markov process. Here we uncover the hidden part of the HMM.
3. Given the observation sequence \mathbf{O} and using the dimensions of the a-priori model, estimate the parameters of a new model $(\theta'_p, \phi'_p, p'_0)$ that maximises the probability in the maximum likelihood sense of seeing the observation sequence \mathbf{O} . Here we re-estimate the parameters of the model using maximum likelihood.

[6 marks]

- (b) Consider that a process can be in one of two states, high (H) or low (L), initially with equal probability. Consider that state H characterises coding DNA while L characterises a non-coding DNA. DNA code is represented by a sequence of four letters A, C, G and T that encode genetic information. Assume that we know the following information.

$$\log_2(P(X_t = H | X_{t-1} = H)) = -1, \quad \log_2(P(X_t = L | X_{t-1} = H)) = -1,$$

$$\log_2(P(X_t = H | X_{t-1} = L)) = -1.3219,$$

$$\log_2(P(X_t = L | X_{t-1} = L)) = -0.7370, \text{ and}$$

$$\log_2(P(Y_t = A | X_t = H)) = -2.3219, \quad \log_2(P(Y_t = A | X_t = L)) = -1.7370,$$

$$\log_2(P(Y_t = C | X_t = H)) = -1.7370, \quad \log_2(P(Y_t = C | X_t = L)) = -2.3219,$$

$$\log_2(P(Y_t = G | X_t = H)) = -1.7370, \quad \log_2(P(Y_t = G | X_t = L)) = -2.3219,$$

$$\log_2(P(Y_t = T | X_t = H)) = -2.3219, \quad \log_2(P(Y_t = T | X_t = L)) = -1.7370.$$

Calculate the most probable state sequence and its \log_2 probability that corresponds to the observation sequence $GGCACTG$ using the Viterbi algorithm.

Please turn over for page 11

Here we need to calculate the \log_2 probabilities but also keep track of the path.

$$\log_2 (P (G, 1)_H) = -2.7370, \quad \log_2 (P (G, 1)_L) = -3.3219, \quad (H)$$

$$\begin{aligned} \log_2 (P (G, 2)_H) &= -1.737 + \max (\log_2 (P (G, 1)_H) + p_{H,H}, \log_2 (P (G, 1)_L) + p_{L,H}) \\ &= -1.7370 + \max (-3.3730, -4.6438) = -5.4740 \quad (HH) \end{aligned}$$

$$\log_2 (P (G, 2)_L) = -6.0589,$$

$$\log_2 (P (C, 3)_H) = -8.2110, \quad (HHH)$$

$$\log_2 (P (C, 3)_L) = -8.7959,$$

$$\log_2 (P (A, 4)_H) = -11.5329,$$

$$\log_2 (P (A, 4)_L) = -10.9480, \quad (HHHL)$$

$$\log_2 (P (C, 5)_H) = -14.0069,$$

$$\log_2 (P (C, 5)_L) = -14.0069, \quad (HHHLL)$$

$$\log_2 (P (T, 6)_H) = -17.3288,$$

$$\log_2 (P (T, 6)_L) = -16.4809, \quad (HHHLLL)$$

$$\log_2 (P (A, 7)_H) = -20.1247,$$

$$\log_2 (P (G, 7)_L) = -18.9549, \quad (HHHLLLL)$$

and so the optimal state sequence is $HHHLLLL$, with \log_2 probability -18.9549 .

[6 marks]

[12 marks]

	Primal (Dual)	Dual (Primal)
	$\max z = \sum_{j=1}^n c_j x_j + z_0$	$\min w = \sum_{i=1}^m y_i b_i + z_0$
1.	$\sum_{j=1}^n a_{ij} x_j = b_i$	$y_i \text{ free}$
2.	$\sum_{j=1}^n a_{ij} x_j \leq b_i$	$y_i \geq 0$
3.	$\sum_{j=1}^n a_{ij} x_j \geq b_i$	$y_i \leq 0$
4.	$x_j \geq 0$	$\sum_{i=1}^m y_i a_{ij} \geq c_j$
5.	$x_j \leq 0$	$\sum_{i=1}^m y_i a_{ij} \leq c_j$
6.	$x_j \text{ free}$	$\sum_{i=1}^m y_i a_{ij} = c_j$