APP MTH 3020 Stochastic Decision Theory

Tutorial 4

Week 9, Friday, October 5

1. Alice owns a collection of rare Mini Coopers and each day has one opportunity to sell it, which she may either accept or reject. The potential sale prices are independently and identically distributed with probability density function $g(x), x \geq 0$. Each day there is a probability $1-\beta$ that the market for Mini Coopers will collapse, making her collection completely worthless.

   (a) There are two states, depending on whether she has sold her collection or not. Denote these states as 0 and 1, respectively. Let $F(1)$ be the value obtained from an optimal policy when she has **not** sold her collection. Write the optimality equation.

   The optimality equation is

$$\begin{aligned} F(1) &= \int_{y=0}^{\infty} \max\{y, \beta F(1)\} g(y) dy \\ &= \beta F(1) + \int_{y=0}^{\infty} \max\{y - \beta F(1), 0\} g(y) dy \\ &= \beta F(1) + \int_{y=\beta F(1)}^{\infty} \{y - \beta F(1)\} g(y) dy. \end{aligned}$$

   (b) Find a policy that maximises her expected return and express it as the unique root of an equation.

   From the optimality equation, we have

$$(1-\beta)F(1) = \int_{y=\beta F(1)}^{\infty} [y - \beta F(1)] g(y) dy. \tag{1}$$

   The left and right hand sides are increasing and decreasing in $F(1)$, respectively, so this equation has a unique root, which we denote as $F(1) = F^*$. Thus, Alice should sell when she can get at least βF^*; that is, we want $y - \beta F(1) = y - \beta F^*$ to be non-negative. Her maximal reward is F^*.

   (b) Show that if $\beta > 1/2$, $g(y) = 2/y^3, y \geq 1$, then Alice should sell the first time the sale price is at least $\sqrt{\beta/(1-\beta)}$.

   If $g(y) = 2/y^3, y \geq 1$, **and** $\beta > 1/2$, then the LHS of (1) is less than the RHS at $F(1) = 1$. Thus, the root is greater than 1, and we compute it as

$$(1-\beta)F(1) = 2/\beta F(1) - \beta F(1)/[\beta F(1)]^2.$$

   Thus, $F^* = 1/\sqrt{\beta(1-\beta)}$, and $\beta F^* = \sqrt{\beta/(1-\beta)}$.

   (c) If $g(y) = 2/y^3, y \geq 1$, and $\beta \leq 1/2$, what should Alice do?

   Sell it at any price!

2. Consider the following infinite-horizon discounted-cost optimality equation for a Markov decision chain with $0 < \beta < 1$, a finite state space $x \in \{1, \ldots, N\}$ and finite action space $a \in \{1, \ldots, M\}$:

$$F(x) = \min_a \left[c(x, a) + \beta \sum_{x_1=1}^{N} F(x_1) \Pr(x_1 | X_0 = x, a_0 = a) \right]. \tag{2}$$

Consider also the linear program

$$\max_{G(1),\ldots,G(N)} \sum_{i=1}^{N} G(i)$$

such that

$$G(x) \leq c(x, a) + \beta \sum_{x_1=1}^{N} G(x_1) \Pr(x_1 | X_0 = x, a_0 = a) \ \text{ for all } \ x, a.$$

This linear program has N variables and $N \times M$ constraints.

(a) Suppose F is a solution to (2). Show that F is a feasible solution to the linear program.
F is feasible as it satisfies all the constrains of the linear program.

(b) Suppose G is also a feasible solution to the linear program. Show that for each x there exists an a such that

$$F(x) - G(x) \geq \beta \mathbb{E}[F(X_1) - G(X_1) | X_0 = x, a_0 = a],$$

and hence that $F \geq G$.

1. If G is a feasible solution, then

$$G(x) \leq c(x, a) + \beta \sum_{x_1=1}^{N} G(x_1) \Pr(x_1 | X_0 = x, a_0 = a) \ \text{ for all } \ x, a.$$

Furthermore, we have

$$F(x) \leq c(x, a) + \beta \sum_{x_1=1}^{N} F(x_1) \Pr(x_1 | X_0 = x, a_0 = a) \ \text{ for all } \ x, a$$

and

$$F(x) = c(x, a) + \beta \sum_{x_1=1}^{N} F(x_1) \Pr(x_1 | X_0 = x, a_0 = a) \quad \text{for some } a \text{ for each } x.$$

Hence,

$$F(x) - G(x) \geq \beta \mathbb{E}[F(X_1) - G(X_1) | X_0 = x, a_0 = a],$$

for some a for each x (consider rearranging the inequalities to get $c(x, a)$, a constant, on one side of the inequalities).

Consequently, as F takes that a at every x, we have $F \geq G$ (for all x).

(c) Finally, argue that F is the unique optimal solution to the LP. What is the use of this result?
The linear program seeks to maximise G; since F is feasible, then $F = G$. Solve infinite-horizon discounted-cost programming problems using linear programming; efficient!

3. A collection of n jobs is to be processed in arbitrary order by a single machine. Job i has processing time p_i and when it completes a reward r_i is obtained, for $i = 1, \dots, n$.

(a) Find the order of processing that maximises the sum of discounted rewards.

Let k represent the time-to-go, i.e. the point at which $(n-k)$th job has just been completed and there remains k uncompleted jobs, say S_k. Let $F_k(S_k)$ be the total discounted reward starting with the k jobs in the set S_k.

The dynamic programming equation is

$$F_k(S_k) = \max_{i \in S_k} \left[r_i \beta^{p_i} + \beta^{p_i} F_{k-1}(S_k - \{i\}) \right].$$

Obviously $F_0(\emptyset) = 0$. Applying the method of dynamic programming we first find $F_1(\{i\}) = r_i \beta^{p_i}$. Then, working backwards, we find

$$F_2(\{i, j\}) = \max \left[r_i \beta^{p_i} + \beta^{p_i + p_j} r_j, r_j \beta^{p_j} + \beta^{p_j + p_i} r_i \right].$$

There will be 2^n equations to evaluate, but continuing in this fashion we can determine $F_n(\{1, 2, \dots, n\})$ as required.

(b) Compare the reward that is obtained by processing the jobs in the order $i_1, \dots, i_k, i, j, i_{k+3}, \dots, i_n$ versus if the order of jobs i and j is reversed. Hence, when is the reward of the first schedule greater than the second? Can you deduce an optimal schedule?

Suppose jobs are ordered $i_1, \dots, i_k, i, j, i_{k+3}, \dots, i_n$. The reward is

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

where $T = \sum_{j=1}^{k} p_{i_j}$, and R_1 and R_2 are, respectively, the sum of the rewards due to jobs coming before and after jobs i, j.

Now, consider the ordering $i_1, \dots, i_k, j, i, i_{k+3}, \dots, i_n$. The reward is

$$R_1 + \beta^{T+p_j} r_j + \beta^{T+p_i+p_j} r_i + R_2.$$

Hence, it is optimal to perform job i before job j if

$$r_i \beta^{p_i} / (1 - \beta^{p_i}) > r_j \beta^{p_j} / (1 - \beta^{p_j}).$$

In fact, it is optimal to schedule jobs in decreasing order of the indices $r_i \beta^{p_i} / (1 - \beta^{p_i})$.