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.
   (b) Find a policy that maximises her expected return and express it as the unique root of an equation.
   (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 $\sqrt{\beta/(1-\beta)}$.
   (c) If $g(y) = 2/y^3, y \geq 1$, and $\beta \leq 1/2$, what should Alice do?

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 $a \in \{1, \ldots, M\}$:

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

   Consider also the linear program

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

   such that

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

   This linear program has N variables and $N \times M$ constraints.
   (a) Suppose F is a solution to (1). Show that F is a feasible solution to 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 $F \geq G$.
   (c) Finally, argue that F is the unique optimal solution to the LP. What is the use of this result?

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 $i = 1, \ldots, n$.
   (a) Find the order of processing that maximises the sum of discounted rewards.
   (b) Compare the reward that is obtained by processing the jobs in the order $i_1, \ldots, i_k, i, j, i_{k+3}, \ldots, i_n$ 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?