Exam Markov Decision Theory and Algorithmic Methods (153192)

January 23, 2007

For each exercise a maximum of 10 points can be obtained.
The maximum total score is 40 points.

1. A student is concerned about her car and does not like dents. When she drives to school, she has a choice of parking it on the street in one space, parking it on the street and taking up two spaces, or parking in the lot. If she parks on the street in one space, her car gets dented with probability 1/10. If she parks on the street and takes two spaces, the probability of a dent is 1/50 and the probability of a \$15 fine is 3/10. Parking in a lot costs \$5, but the car will not get dented. If her car gets dented, she can have it repaired, in which case it is out of commission for one day and costs her \$50 in fees and cab fares. She can also drive her car dented, but she feels that the resulting loss of value and pride is equivalent to a cost of \$9 per school day. She wishes to determine the optimal policy for where to park and whether to repair the car when dented in order to minimise her (long-run) expected average cost per school day.
   (a) Formulate this problem as a Markov decision problem by indentifying the states, actions, immediate rewards and transition probabilities.
   (b) Carry out one iteration of the policy iteration algorithm, starting with the decision rule in which the student always parks on the street in one space and is not bothered about getting dents repaired. Interpret the new decision rule.
   (c) Show that it is optimal to park in one space if the car is not dented and to have the car repaired if it is dented. Also determine the minimal expected average cost.
   (d) Formulate the dual linear program that corresponds to this Markov decision problem. (You need *not* solve it.)
   (e) How can you determine an optimal strategy from the optimal solution of the dual linear program?

2. Consider a Markov decison problem in which the expected total discounted reward is maximised. Future rewards are discounted with factor λ per period. Let $v^0(s)$ be a uniform bounded function, that is $|v^0(s)| < M$, $s \in S$, for some $M > 0$. For $n = 0, 1, 2, \ldots$ define the function v^{n+1} by recursion (in vector notation)

$$v^{n+1} = r_d + \lambda P_d v^n$$

   for a given randomized Markovian decision rule d.

(a) Show that the sequence of functions v^{n+1} has the following property: if $v^{n-1}(s) \leq v^n(s)$ for all $s \in S$ then also $v^n(s) \leq v^{n+1}(s)$ for all $s \in S$.

(b) Show that $\lim_{n\to\infty} v^n(s) = v_\lambda^{d^\infty}(s)$, where $v_\lambda^{d^\infty}$ is the expected total discounted reward of policy d^∞.

(c) Show that in the policy improvement step of the policy improvement algorithm the new decision rule d_{n+1} satisfies

$$r_{d_{n+1}} + \lambda P_{d_{n+1}} v_\lambda^{(d_n)^\infty} \geq v_\lambda^{(d_n)^\infty}, \tag{1}$$

where the inequality is componentwise.

(d) Assume that the decision rule d_{n+1} in (c) is such that the inequality (1) is strict for at least one state $s = s_0$. Show that $v_\lambda^{(d_{n+1})^\infty}(s) \geq v_\lambda^{(d_n)^\infty}(s)$ for all states s and $v_\lambda^{(d_{n+1})^\infty}(s_0) > v_\lambda^{(d_n)^\infty}(s_0)$. (Use the results from (a) and (b).)

3. Consider the symmetrical shortest queue problem handling customers with Poisson arrivals with arrival rate λ and exponential service requests with mean $1/\mu$. Let $N = \{N(t),\ t \geq 0\}$ record the number of customers at the two queues.

(a) Show that N is a Markov chain, and give the state space, and the transition rates.

(b) Modify N to the Markov chain M recording the length of the shortest queue, and the difference in queue lengths. Give the transition diagram for M (picture with transitions and corresponding rates).

(c) Give the global balance equations for M.

(d) Let $p(m,n) = \alpha^m \beta^n$. Characterise all α and β for which p satisfies the global balance equations at the interior of the state space for M, that is for $\{(m,n) : m > 0, n > 0\}$.

(e) Select α, β that intuitively best characterises the equilibrium distribution for states far from $(m,n) = (0,0)$, and provide the intuitive argument for this selection.

(f) Derive rates at the boundaries $m = 0, n > 0$, and $m > 0, n = 0$ such that the distribution $p(m,n) = \alpha^m \beta^n$ selected above is the unique equilibrium distribution. (If you did not find intuitively best choice above, then just take a solution for α and β for which p satisfies the global balance equations at the interior of the state space.)

4. Consider the single server queue with Markov Modulated Arrivals. The arrival rate is modulated by a four state Markov chain X, with states $\{1,2,3,4\}$, and transitions rates s_{ij} from states i to j equal to $s_{12} = s_{23} = s_{34} = s_{41} = \alpha$, and $s_{21} = s_{32} = s_{43} = \beta$. Other transitions among the states of X are not allowed. In states $1,2$ of X, the arrival rate to the single server queue is λ_1, and in states $3,4$ the arrival rate is λ_2. Customers arriving to the single server queue all have exponential service requirement with mean $1/\mu$.

(a) Give the diagram with transitions and transition rates for this queue, and specify the levels and phases.

(b) Provide a description in matrix geometric form, that is specify the matrices A_0, A_1, A_2.

(c) Consider the discrete time single server queue with Markov Modulated Arrivals obtained from the queue described above via uniformization. Give the transition probabilities for this queue.

(d) Describe this discrete time queue in matrix geometric form, i.e. specify the matrices describing the transitions between levels, and within levels. Specify the matrices B_0, B_1, B_2 for the transitions between and inside levels.

(e) Derive an equation for the equilibrium distribution in matrix geometric form, i.e., $\pi_n = \pi_0 R^n$: give an equation for the matrix R, and an equation for π_0.

(f) Obtain an explicit expression for R, that is solve the matrix equation for R.

(g) We are now interested in the phase of the discrete time matrix geometric queue. This phase is obtained as the equilibrium distribution of the Markov chain with transition probability matrix $B = B_0 + B_1 + B_2$. Obtain this equilibrium distribution via Gaussian elimination.