Exam Markov Decision Theory and Algorithmic Methods (153192)

April 11, 2008

This exam consists of three exercises with the following weights:

exercise	1	2	3
points	20	10	10

Plan your time. Motivate all your answers.
When a derivation is required, you most provide the derivation. *Good luck!*

1. **Tandem queues with loss**
Consider a tandem network model of two servers. New customers arrive according to a Poisson process with rate λ and join a ***finite*** waiting room in front of server 1. They receive service, exponentially distributed with rate μ_1, then they move to a second waiting room and eventually receive a second service, exponentially distributed with rate μ_2. The waiting room between the first and second server has ***infinite*** capacity, however the waiting room in front of server 1 has finite capacity and may contain (including the customer in service) at most C customers. When the waiting is full, arrival customers are lost. We will model the system in matrix geometric form.

   (a) Give the transition diagram with transition rates of this model.

   (b) Specify the levels and phases of this model, and define the matrices B, A_0, A_1, A_2 of QBD process.

   (c) Give the stability condition in terms of the matrices A_0, A_1, A_2. The objective now is to write the stability condition explicitly in terms of λ, μ_1, μ_2, C (Note that questions c.i and c.ii are independent of the rest of questions):

      i. Find the equilibrium distribution of the irreducible Markov process of generator $A=A_0+A_1+A_2$.

      ii. Derive the stability condition in terms λ, μ_1, μ_2, C .

   (d) Prove that **R**, QBD rate matrix, ***cannot*** be solved in closed form via the balance principle. Propose an iterative approach to find **R**.

   (e) Specify the equilibrium distribution of level 0 denoted as Π_0 and give the expression for the equilibrium distribution.

   (f) Compute the probability of customer loss.

2. 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?

3. 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).)