Exam Markov Decision Theory and Algorithmic Methods (191531920)

January 27, 2012 8:45-11:45 hrs

This exam consists of 6 exercises.
Motivate all your answers.

1. Consider a worker who begins each period with a current wage offer and has two actions. He can work at that wage or he can search for a new wage offer. If he chooses to search, he earns nothing during the current period, and his new wage is drawn according to some probability measure f from the interval $W = [0, \overline{w}]$. He cannot divide his time within a period between searching and working. Moreover, if a worker chooses to work during the period, then with probability $1-\theta$ the same wage is available to him next period. But with probability θ he will lose his job at the beginning of next period and begin next period with a "wage" of zero.

   The worker does not value leisure. Let $0 < \beta < 1$ be the discount factor, and $U : \mathbb{R}_+ \to \mathbb{R}_+$ the worker's utility function, which is continuously differentiable, strictly increasing and strictly concave with $U(0) = 0$ and $U'(0) < \infty$. If he earns c_t during period t then his total expected utility equals $E[\sum_{t=0}^{\infty} \beta^t U(c_t)]$. Let v be the supremum of this expected discounted utility.

   (a) Show that v must satisfy

   $$v(w) = \max\left\{U(w) + \beta[(1-\theta)v(w) + \theta v(0)], \beta \int_0^{\overline{w}} v(w')f(w')dw'\right\}.$$

   (b) Define $A = \beta \int_0^{\overline{w}} v(w')f(w')dw'$. Note that $v(0) = A$. Show that there is a unique $w^* \in W$ such that

   $$v(w^*) = U(w^*) + \beta[(1-\theta)v(w^*) + \theta A] = A.$$

   (c) Show that v has the form

   $$v(w) = \begin{cases} A & \text{if } w < w^*, \\ \frac{U(w)+\beta\theta A}{1-\beta(1-\theta)} & \text{if } w \geq w^*. \end{cases}$$

   (d) What does the optimal decision rule for the worker look like? Hint: it depends on w^*.

2. Consider the following infinite-horizon average reward Markov Decision Problem (MDP). There are two states, $S = \{s_1, s_2\}$. In state s_1 the available actions are a_1 and a_2, with $r(s_1, a_1) = 5$, $r(s_1, a_2) = 10$, $p(s_2|s_1, a_1) = 1/2$, $p(s_2|s_1, a_2) = 1$. In state 2 there is one action a_3 with $r(s_2, a_3) = -2$, and $p(s_2|s_2, a_3) = 1/2$.

   (a) What are the optimality equations for this particular MDP?

   (b) Determine the optimal gain and average optimal stationary policy.

3. Consider an infinite horizon discounted MDP with discount factor λ.

   (a) Show that $|r(s, a)| \leq M$ implies $\|v_\lambda^*\| \leq M/(1-\lambda)$.

   (b) What are the optimality equations for this MDP?

   (c) Define $\mathcal{L}v = \sup_{d \in D^{\text{MD}}}(r_d + \lambda P_d v)$. Let $v \in V$. Prove that if $v \leq \mathcal{L}v$ then $v \leq v_\lambda^*$.

4. One way to find the equilibrium distribution of a large, aperiodic discrete-time Markov chain $\{X_n, n = 0, 1, 2, \ldots\}$ is to determine powers P^k of the transition matrix P.

   (a) Give upper and lower bounds on $p_j = \lim_{t\to\infty} P(X_n = j)$ based on the matrix P^k, and prove for the lower bound that it is nondecreasing in k, and therefore indeed a lower bound on p_j.

   (b) Explain in the standard power method (iterating $p^{(n+1)} = p^{(n)}P$ until $p^{(n+1)} - p^{(n)}$ is small) why this scheme converges geometrically to the steady-state vector p and give the decay rate (i.e. rate of convergence).

5. Let $\{X(t), t \geq 0\}$ be a birth-death process on $\{0, 1, 2\}$ with birth rates $\lambda_0 = 1$ and $\lambda_1 = 2$, and death rates $\mu_1 = \mu_2 = 3$.

   (a) Use uniformization to define a discrete time Markov chain $\{Y_n\}$ that has the same stationary distribution as $\{X(t)\}$. In particular give its transition matrix P.

   (b) Use it to give an expression for $p_{02}(4) = P(X(4) = 2|X(0) = 0)$, and indicate how this can be used numerically.

6. Consider any quasi-birth-death process in continuous time, with m phases, where the transitions from states in level 0 to level 1 are the same as those from states in level i to level $i + 1$, $i > 0$.

   (a) Write down the general structural form of the infinitesimal generator Q, and give the balance equations for the row vectors $p_i, i = 0, 1, 2, \ldots$; here $p_i = (p_{i0}, p_{i1}, \ldots, p_{i,m-1})$ contains the equilibrium probabilities for the states in level i.

   (b) Under which condition is the process positive recurrent?

   (c) If this condition holds, the solution can be written as $p_i = p_0 R^i$. How can p_0 and R be found in general?

   Now consider an $M/M/1$ queue with arrival rate $\lambda = 1$ and service rate $\mu = 2$, with the additional feature that the server turns 'off' as soon as it becomes idle, and that it needs to start up (turn 'on') when a customer arrives to an empty system. During start-up, which takes an exponentially distributed amount of time with mean γ^{-1}, no customers are served. γ = 3

   (d) Solve the equilibrium distribution *by using the spectral expansion method.* Hint: start out by trying solutions of the form $p_i = yx^i, i = 0, 1, 2, \ldots$.

   (e) Give the steady state mean number of customers in the system.

<u>Norm:</u>

1				2		3			4		5		6					total
a	b	c	d	a	b	a	b	c	a	b	a	b	a	b	c	d	e	
2	3	1	1	2	4	3	1	1	3	2	2	2	2	1	2	3	1	+ 4 = 40