Practice Problem Set 4

IE 708 Markov Decision Processes
IEOR, IIT Bombay

Monsoon Sem '24

Exercise 1 (Machine replacement) A machine can be in states $S = \{1, 2, \ldots, n\}$, with higher numbered states denoting higher machine deterioration and operating cost. The income in state i for one period is $c(i)$, which is non-increasing in i. At each period, there are two options: (a) let the machine operate for one more period in the same state, or (b) repair the machine with a cost of R and bring it to state 1 (corresponding to the state of the machine with the least operating cost). In option (a) the transition probabilities p_{ij}, are such that $p_{ij} = 0$ for $j < i$.

(a) Find the partitions S_c, S_t of S and verify that blueWeak Accessibility (WA) or weakly communicating condition holds.

(b) Write the optimality equation for the average cost.

(c) Are all stationary policies uni-chain? If not, give an example stating the two (or more) recurrent classes.

Exercise 2 (PI is not bias optimal in lex ordering sense) Let $S = \{s_1, s_2\}$, $A_{s_1} = \{a_{1,1}, a_{1,2}\}$ and $A_{s_2} = \{a_{2,1}\}$; $r(s_1, a_{1,1}) = 4, r(s_2, a_{2,1}) = 8$; and $p(s_1|s_1, a_{1,1}) = 1, p(s_2|s_1, a_{1,2}) = 1, p(s_1|s_2, a_{2,1}) = 1$.

(a) Start PI algo with decision rule $\delta(s_1) = a_{1,2}$. In policy evaluation step, compute h using $P^*_{d_0} h = 0$.

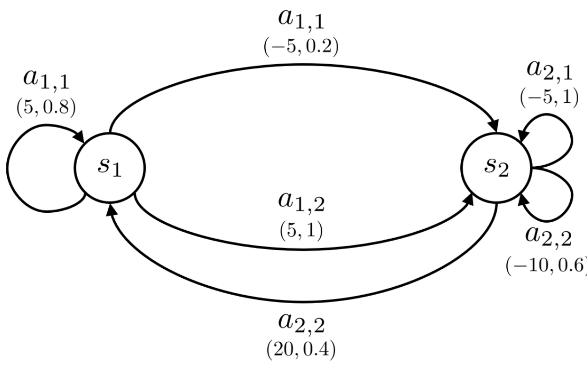
(b) Start PI algo with decision rule $\gamma(s_1) = a_{1,1}$. In policy evaluation step, compute h using $P^*_{d_0} h = 0$.

(c) Compare the bias vector in (a) and (b).

Exercise 3 (a) Write the primal LP for the model in Figure 3, for infinite horizon average expected reward.

(b) Write the Dual LP for the above.

(c) Identify the optimal decision rule based on the solution of the above LPs. (Which LP would you use?)



(d) What do the variables of dual LP represent?

(e) Under the optimal decision rule identified in (c), what is the steady state probability that system is in s_1.

(f) Add a constraint of the form $\sum_{s \in S} \sum_{a \in A_s} c(s,a)x(s,a) \leq C$ such that under the optimal decision rule the steady state probability that system is in s_1 is no more than α.

(g) Find the optimal solution for $\alpha = 0.5$

(h) For what value of α will the solution be the same as that of the original LP without the additional constraint from (f).

Exercise 4 Write the primal and the dual LP for the model in Figure 3, the discounted case. Solve with discount factor 0.9.

Exercise 5 Consider one of the classical MDP models, *admission control to a queue*. First, this *queue or congestion model* can be any *service facility*.

Arrivals to this facility are a Poisson process with rate λ. These arrivals, also called customers, seek service from the service facility; the service facility offers service to each one of them in a sequential manner, as long as at least one of them is waiting for service. Let the service times of these customers be independent and identically exponentially distributed with rate μ. And this service process be independent of the arrival process so that $\cdots$.

- Check that the stochastic processes of the number of customers, usually called jobs, is a continuous time Markov chain with state space $\{0, 1, 2, \cdots$. Btw, such a CTMC is also called birth-death process; why?

- Consider now a *control* version of this problem.

  The queue manager charges each customer ₹R as service charges while incurs ₹h per unit time for each waiting customer to provide *waiting space* to the customers (jobs) waiting to avail service.

  Do you see a trade-off between revenue generated vs cost incurred by the queue manager?

- A way to handle the above trade-off is this: The queue manager can either admit or don't admit each arrival customer (job). Of course, that (lucky) customer arriving to an empty system (when the number in the system is zero) is always admitted.

- Let money be discounted at rate α per unit time.

- Port the above continuous time MDP as a discrete time MDP, without loss of optimality. (*uniformization*, Richard Serfozo, Steve Lippmann $\cdots$).