Problem Set 6

CS 6375

Due: 12/9/2018 by 11:59pm

Note: all answers should be accompanied by explanations for full credit. Late homeworks will not be accepted.

Problem 1: Markov Decision Processes (35pts)

1. Consider a Markov decision process with four states s_1, s_2, s_3, s_4 and four actions a_1, a_2, a_3, a_4 with the following transition and reward functions.

   - $T(s_i, a_j) = s_j$ for all $i, j \in \{1, 2, 3, 4\}$.
   - $R(1, 2) = R(2, 3) = 1$, $R(3, 1) = R(4, 1) = -1$, the rewards of the form $R(s, s) = 0$ for all states s, and the rewards for the remaining transitions are equal to .5.

   Using the above MDP, answer the following questions.

   (a) How many possible deterministic policies are there for this MDP?

   (b) For $\gamma = .8$, find the optimal value function V^* and an optimal policy π^*. Is there a unique optimal deterministic policy?

   (c) If the policy in part (b) is not required to be deterministic, is the optimal stochastic policy unique?

   (d) How does the optimal policy change if $\gamma = .01$?

2. For any MDP and any two policies π_1 and π_2, show that there exists a policy π_3 such that $V^{\pi_3}(s) \geq V^{\pi_1}(s)$ for all $s \in S$ and $V^{\pi_3}(s) \geq V^{\pi_2}(s)$ for all $s \in S$.

3. As we saw in class, Markov decision processes make decisions based on the current state of the environment and a chosen policy. Suppose that you are given an MDP, but you would like the agent's decision to depend on the last two states of the environment instead of just the last. Can this requirement be formulated in the MDP framework? Explain why or why not.

Problem 2: Poisson Maximum Likelihood Estimation (30pts)

Consider a nonnegative, integer-valued random variable X that is distributed according to a Poisson distribution $X \sim \frac{\lambda^x e^{-\lambda}}{x!}$ for some real-valued parameter $\lambda > 0$.

1. Given data samples $x^{(1)}, \ldots, x^{(m)}$, what is the maximum likelihood estimate for λ?

2. Suppose you are interested in bounding the sample complexity. Using the Chernoff bound below, derive a lower bound on the number of samples needed to guarantee that $\lambda_{MLE} \leq \lambda + \epsilon$ for some $\epsilon > 0$ with probability at least $1 - \exp(-5)$. For full credit, your bound should be the best possible.

   Chernoff bound: for a random variable Y and a real number $t > 0$,

   $$\Pr(Y \geq a) \leq \frac{\mathrm{E}(\exp(tY))}{\exp(ta)}$$

   where $E(\exp(tY))$ is the expected value of $\exp(tY)$.

   Hint: $\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$.

3. Suppose now that you introduce a prior probability distribution, $\lambda \sim \frac{1}{5}\max\{-\lambda/10 + 1, 0\}$. What is the MAP estimate under this prior probability distribution?

Problem 3: Neural Networks (35pts)

Using only perceptrons, construct a neural network for the following input/output pair: Takes 10 binary inputs with one binary output which is 1 if the number of ones in the input is divisible by four and zero otherwise. Generate a data set for this problem and use MATLAB or Python to learn the weights of a relu (e.g., $\max\{0, \cdot\}$ or its smooth analog) neural network with the same structure as you produced for the first part of this question. You should use a cross entropy loss. How do the learned weights compare to your perceptron solution as you vary the size of the training set?

Course Evaluation:

If you haven't done so already, please go to `eval.utdallas.edu` and provide feedback on the course.