CS 4758/6758: Robot Learning: Homework 4

Due: Apr 14 5pm

1 Logistic Regression

We have $x \in \Re^2$ and $y \in \{0, 1\}$.

(a) We now have $M = 4$ points: (-1,-1) and (-1,1) labeled as $y = 0$, and (1,-1) and (1,1) labeled as $y = 1$. Where would the decision boundary for the logistic classifer be? Please draw.

(b) Now we have add another point (1,0) labeled as $y = 1$, thus having $M = 5$ total points. Would the decision boundary change? Would it change at all? Justify.

(c) Now we add another point (100,0) labeled as $y = 1$, thus having $M = 6$ total points. Would the decision boundary change? Would it change at all? Justify.

(d) One student was lazy: instead of running the logistic classifier for the 3 parts above, he ran a linear regression. I.e., he considered $y \in \Re$. Please redo the three parts above. And comment on what would happen.

2 Kernels

Prove that Radial Basis Function is a Kernel.

3 Reinforcement Learning

In this problem, you have to show that MDP is guaranteed to find the optimal policy. Consider an MDP with finite state and action spaces, and discount factor. Let B be the Bellman update operator with V a vector of values for each state. I.e., if $V = B(V)$, then

$$V'(s) = R(s) + \gamma \max_{a \in A} \sum_{s' \in S} P_{sa}(s') V(s')$$

(a) Prove that if $V_1(s) \leq V_2(s)$ for all $s \in S$, then $B(V_1)(s) \leq B(V_2)(s)$ for all $s \in S$

(b) Let B^π be the Bellman operator with respect to a particular policy π. Prove that for any V,

$$||B^\pi(V) - V^\pi||_\infty \leq \gamma ||V - V^\pi||_\infty$$

where $||V||_\infty = max_{s\in S}|V(s)|$.

Intuitively, this means that applying the Bellman operator B to any value function V, brings that value function closer to the value function for π, V^π. This also means that applying B^π repeatedly (an infinite number of times)

$$B^\pi(B^\pi(B^\pi \cdots B^\pi(V)\cdots))$$

will result in the value function V^π (a little bit more is needed to make this completely formal, but we will not worry about that here). Use the fact that for any $\alpha, x \in R^n$, if $\sum_i a_i = 1$ and $a_i > 0$, then $\sum_i \alpha_i x_i \leq \max_i x_i$.

(c) We say that V is a fixed point of B^π if $B^\pi(V) = V$. Using the fact that the Bellman update operator is a γ-contraction in the max-norm, prove that B^π has at most one fixed point -i.e., that there is at most one solution to the Bellman equations. You may assume that B^π has at least one fixed point.

(d) Now suppose that we have some policy π, and use Policy Iteration to choose a new policy π' according to

$$\pi'(s) = \arg\max_{\alpha\in A} \sum_{s'\in S} P_{sa}(s')V^\pi(s')$$

Show that this policy will never perform worse that the previous one, i.e., show that for all $s \in S$; $V^\pi(s) \leq V^{\pi'}(s)$. In order to show it, first show that $V^\pi(s) \leq B^{\pi'}(V^\pi)(s)$, then use the a) and b) to show that $B^{\pi'}(V^\pi)(s) \leq V^{\pi'}(s)$.

4 Kalman Filters

4.1 Multiple sensors

Suppose we have a continuous system with state $x \in \Re^N$. The state at time k is determined by the equation:

$$x_{k+1} = x_k + w_{k+1} \quad \text{where } w_{k+1} \sim N(0, Q)$$

Now we have two sensors z_k and y_k. One sensor gives data at $k = 1, 3, 5, \ldots$, and the other sensor gives data at $k = 2, 4, 6, \ldots$.

$$z_k = x_k + v_k \quad \text{where } v_k \sim N(0, R_1)$$
$$y_k = x_k + e_k \quad \text{where } e_k \sim N(0, R_2)$$

Derive the time and measurement update equations for this situation.

4.2 Kalman filter Matlab implementation

Unfortunately, you've been assigned to work on an old robot with unreliable sensors. You are particularly interested in measuring the room temperature in which the robot operates in. Therefore, you decide to use a Kalman filter.

Let x_k denote the room temperature at time k and z_k the sensor measurement at time k.
Since the robot will be operating in the same room then you don't expect the temperature to change. This translates to the following system dynamics:

$x_k = x_{k-1} + w_k$ where $w_k \sim N(0, Q)$

$z_k = x_k + v_k$ where $v_k \sim N(0, R)$

Keep in mind that room temperature is between 20 and 25 C (use this to initialize x_0 and P_0[1] accordingly). Furthermore, the temperature sensor's error variance is 4. Finally, you believe that $0.00001 < Q < 0.0001$.
Run a Kalman filter on the measurement data in measurements.txt. Plot the measurement data and the Kalman filter's state estimate against time. Also report the values of x_0, Q, and R which you used as well as the mean and standard deviation of the state estimate over the last 20 steps. Attach your code.

[1] The a posteriori estimate error covariance at time 0