Homework #3
Likelihood Estimation & Linear classifier
Statistical Pattern Recognition

Due Azar 14th,1395
Format of homework file: Archive all files in a folder named as your student number and send it to mohammadhme@gmail.com. Send your emails with a subject of PR95F3_xxxxxx (replace xxxxxx by your student number)

---

1. Let $\{x_k\}, k = 1,2,\dots,N$ denote independent training from one of the following densities. Obtain the Maximum Likelihood estimate of θ in each case.
   a. $f(x_k;\theta) = \frac{x_k}{\theta^2}\exp\left(-\frac{x_k^2}{2\theta^2}\right)$ $\quad x_k \geq 0 \quad \theta > 0 \quad$ Rayleigh Density
   b. $f(x_k;\theta) = \sqrt{\theta}x_k^{\sqrt{\theta}-1}$ $\quad 0 \leq x_k \leq 1 \quad \theta > 0 \quad$ Beta Density

---

2. Let x have uniform density

$$f_x(x|\theta) \sim U(0,\theta) = \begin{cases} \frac{1}{\theta} & 0 \leq x \leq \theta \\ 0 & otherwise \end{cases}$$

   a. Suppose that n samples $D = \{x_1, x_2, \dots, x_n\}$ are drawn independently according to $f_x(x|\theta)$. Show that the maximum likelihood estimate for θ is $\max[D]$, i.e., the value of the maximum element in D.
   b. Suppose that $n = 5$ point are drawn from the distribution and the maximum value of which happens to be $\max_k x_k = 0.6$. Plot the likelihood function $f_x(D|\theta)$ in the range $0 \leq \theta \leq 1$. Explain in words why you do not need to know the values of other four points.

---

3. Consider the standard two class SVM with the hinge loss. Argue that under a given value of regularization parameter:

$$\text{Leave-one-out Error} < \frac{\#SV_s}{l}$$

   Where l is the size of training data and $\#SV_s$ is the number of support vectors obtained by training SVM on the entire set of training data.

---

4. Consider the 2-dimensional points and their classification ('+' or '-') below:

x	y	class
0	4	+
8	3	+
6	-2	-
4	0	-
2	1	-

   a. The points with classification '+' and '-' corresponds to the point sets M_+ and M_-, respectively. Draw the points and determine first whether or not the sets M_+ and M_- are linearly separable. And then whether or not the two sets are linearly separable by a 2-dimensional perceptron.

b. Manually execute the perceptron learning algorithm on this dataset. Based on your answer from part (a) decide whether or not you need a bias. Use a vector of all ones as the initial weight vector. Write all the intermediate results of your perceptron computation in a table.
c. Give the linear function that has been learned by this perceptron.
d. Classify point (5,2) base on the trained perceptron.

---

5. [Computer Project] Implement the following projects from the reference book.

   Duda, Richard O., Peter E. Hart, and David G. Stork. Pattern classification. John Wiley & Sons, 2012

   Chapter 5: Computer Exercises 1,2 and 9