

BIOMETRICS
CSE190
Winter 2011

Assignment 2
Due: February 10, 2011

1. Duda, Hart, Stork 3.1.

Let x have an exponential density

$$p(x|\theta) = \begin{cases} \theta e^{-\theta x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

- Plot $p(x|\theta)$ versus x for $\theta=1$
- Plot $p(x|\theta)$ versus θ over $(0 \leq \theta \leq 5)$ for $x=2$.
- Suppose that n samples $x_1, \dots, x_n$ are drawn independently according to $p(x|\theta)$. Show that the Maximum-likelihood estimate for θ is given by:

$$\hat{\theta} = \frac{1}{\frac{1}{n} \sum_{k=1}^n x_k}$$

2. Duda, Hart, Stork 3.35.

Let the sample mean μ_n and the sample covariance matrix C_n for a set of n samples $\mathbf{x}_1, \dots, \mathbf{x}_n$ (each of which is d -dimensional) be defined by

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i$$
$$C_n = \frac{1}{n-1} \sum (\mathbf{x}_i - \mu_n)(\mathbf{x}_i - \mu_n)^t$$

We call these the “nonrecursive” formulae.

- What is the computational complexity of calculating μ_n and C_n by these formulae?
- Show that the alternative” recursive techniques based on successive addition of new samples $\mathbf{x}_{n+1}$ can be derived using the recursive relations

$$\mu_{n+1} = \mu_n + \frac{1}{n+1}(\mathbf{x}_{n+1} - \mu_n)$$
$$C_n = \frac{n-1}{n} C_n + \frac{1}{n+1}(\mathbf{x}_{n+1} - \mu_n)(\mathbf{x}_{n+1} - \mu_n)^t$$

- What is the computational complexity of finding μ_n and C_n by these recursive methods?
- Describe situations where you might prefer to use the recursive method for computing μ_n and C_n , and ones where you might prefer the nonrecursive method?

3. Consider a normal $p(x)=N(\mu, \sigma^2)$ and Parzen window function $\phi(x) = N(\mu, 1)$. Show that the Parzen window estimate

$$p_n(x) = \frac{1}{nh_n} \sum_{i=1}^n \phi\left(\frac{x - x_i}{h_n}\right)$$

has the following property

$$E[p_n(x)] = N(\mu, \sigma^2 + h_n^2)$$

4. Consider the following set of two dimensional vectors from three categories:

ω_1		ω_2		ω_3	
X_1	X_2	X_1	X_2	X_1	X_2
10	0	5	10	2	8
0	-10	0	5	-5	2
5	-2	5	5	10	-4

- Plot the decision boundary resulting from the nearest neighbor rule just for categorizing ω_1 and ω_2 . Find the sample mean $\mathbf{m}_1$ and $\mathbf{m}_2$ and on the same figure sketch the decision boundary corresponding to classifying x by assigning it to the category of the nearest sample mean.
- Repeat part (a) for categorizing only ω_1 and ω_3 .
- Repeat part (a) for categorizing only ω_2 and ω_3 .
- Repeat part (a) for three-category classifier, classifying ω_1 , ω_2 and ω_3 .