# Machine Learning and Computational Statistics

**Theodoros Georgiopoulos**

Homework 6

## Exercises for Unit 6: Parametric pdf estimation

### Solution for exercise 1

We have the problem

$$\min J(\mu) = \sum_{n=1}^{N} (x_n - \mu)^2$$
$$\text{subject to } (\mu - \mu_0)^2 \leq \rho.$$

We define the Lagrangian function

$$L(\mu) = \sum_{n=1}^{N} (x_n - \mu)^2 + \lambda((\mu - \mu_0)^2 - \rho).$$

The derivative with respect to $\mu$ is

$$\begin{aligned}
\frac{\partial L(\mu)}{\partial \mu} &= -2 \sum_{n=1}^{N} (x_n - \mu) + 2\lambda\mu - 2\lambda\mu_0 \\
&= -2 \sum_{n=1}^{N} x_n + 2N\mu + 2\lambda\mu - 2\lambda\mu_0.
\end{aligned}$$

We set the derivative equal to zero and we solve the equation for finding $\mu$.

$$\frac{\partial L(\mu)}{\partial \mu} = 0 \iff \mu = \frac{\lambda\mu_0 + \sum_{n=1}^{N} x_n}{N + \lambda}$$

### Solution for exercise 2

The data are modeled by a pdf of the form

$$p(\boldsymbol{x}) = \sum_{j=1}^{m} P_j p(\boldsymbol{x}|j), \quad \sum_{j=1}^{m} P_j = 1, \quad \int_{-\infty}^{+\infty} p(\boldsymbol{x}|j) = 1$$

The parameter updating part of the Expectation Maximixation (EM) algorithm is

$$\arg\ max_{P_1, P_2, \ldots, P_M} \sum_{i=1}^{N} \sum_{j=1}^{m} P(j|\boldsymbol{x_i}) \ln P_j$$
$$\text{subject to} \sum_{j=1}^{m} P_j = 1.$$

(a) We define the Lagrangian function

$$L(P_1, P_2, \ldots, P_m) = \sum_{i=1}^{N} \sum_{j=1}^{m} P(j|\boldsymbol{x_i}) \ln P_j + \lambda(\sum_{j=1}^{m} P_j - 1).$$

(b) We set the derivative of the Lagrangian function with respect to $P_j$ equal to zero and we solve for $P_j$.

$$\frac{\partial L(P_1, P_2, \ldots, P_m)}{\partial P_j} = 0 \iff \sum_{i=1}^{N} P(j|\boldsymbol{x_i}) \frac{1}{P_j} + \lambda = 0$$

So,

$$P_j = -\frac{\sum_{i=1}^{N} P(j|\boldsymbol{x_i})}{\lambda} \tag{1}$$

(c) Also we have

$$\sum_{j=1}^{m} P_j = 1. \tag{2}$$

From (1),(2)

$$\sum_{j=1}^{m} \frac{\sum_{i=1}^{N} P(j|\boldsymbol{x_i})}{\lambda} = -1$$

and

$$\lambda = -\sum_{j=1}^{m} \sum_{i=1}^{N} P(j|\boldsymbol{x_i}) = -N. \tag{3}$$

(d) Finally, from (1),(3)

$$P_j = \frac{\sum_{i=1}^{N} P(j|\boldsymbol{x_i})}{N}.$$