Homework Assignment #2

Assigned: Thursday 02/19/2015; Due: Thursday 03/05/2015 at 11:59pm via Oncourse.
(total: 110 points)

Problem 1 (15 points) Suppose that the number of accidents occurring daily in a certain plant has a Poisson distribution with an unknown mean λ. Based on previous experience in similar industrial plants, suppose that our initial feelings about the possible value of λ can be expressed by an exponential distribution with parameter $\theta = \frac{1}{2}$. That is, the prior density is

$$f(\lambda) = \theta e^{-\theta\lambda}$$

where $\lambda \in (0, \infty)$. If there are 79 accidents over the next 9 days, determine

a) (5 points) the maximum likelihood estimate of λ.
b) (5 points) the maximum a posteriori estimate of λ.
c) (5 points) the Bayes estimate of λ.

Problem 2 (15 points) Let $X_1, \ldots, X_n$ be i.i.d. Gaussian random variables, each having an unknown mean θ and known variance σ_0^2. If θ is itself selected from a normal population having a known mean μ and a known variance σ^2

a) (5 points) what is the maximum a posteriori estimate of θ?
b) (10 points) what is the Bayes estimate of θ?

Problem 3 (20 points) Let $X_1, \ldots, X_n$ be i.i.d. random variables with probability distribution

$$f(x|\alpha) = \alpha^x (1-\alpha)^{1-x}$$

where $x \in \{0,1\}$. Assuming that the unknown parameter α that was selected from a (0,1) uniform distribution find the Bayes estimator of α.

Problem 4 (20 points) Consider the following minimization problem:

$$\underset{\mathbf{x}}{\operatorname{argmin}} \|\mathbf{Ax} - \mathbf{b}\|$$

where $\mathbf{A}$ is a m-by- n matrix, $\mathbf{x}$ is a n-by- 1 vector and $\mathbf{b}$ is a m-by- 1 vector (all vectors and matrices are real). Owing to the fact that the row space and nullspace of $\mathbf{A}$ are orthogonal, any vector $\mathbf{x} \in \mathbb{R}^n$ can be decomposed as $\mathbf{x} = \mathbf{x}_r + \mathbf{x}_n$, where $\mathbf{x}_r$ lies in the row space of $\mathbf{A}$ and $\mathbf{x}_n$ lies in the nullspace of $\mathbf{A}$. Suppose now that $\hat{\mathbf{x}} = \hat{\mathbf{x}}_r + \hat{\mathbf{x}}_n$ is one solution to the minimization problem above.

a) (5 points) Prove that $\hat{\mathbf{x}} = \hat{\mathbf{x}}_r + \alpha \hat{\mathbf{x}}_n$, where $\alpha \in \mathbb{R}$, is also a solution to the minimization problem.
b) (15 points) Prove that $\hat{\mathbf{x}}_r$ from above is common to all solutions that minimize $\|\mathbf{Ax} - \mathbf{b}\|$. In other words, prove that there is no other vector from the row space that can be combined with any vector from the nullspace to minimize $\|\mathbf{Ax} - \mathbf{b}\|$.

Problem 5 (20 points) Expectation-Maximization. Let X be a random variable distributed according to $p_X(x)$ and Y be a random variable distributed according to $p_Y(y)$. Let $D_X = \{x_i\}_{i=1}^m$ be an i.i.d. sample from $p_X(x)$ and $D_Y = \{y_i\}_{i=1}^n$ be an i.i.d. sample from $p_Y(y)$. Let $D = D_X \cup D_Y$. Furthermore, define $p_X(x)$ and $p_Y(y)$ as follows:

$$p_X(x) = \alpha N(\mu_1, \sigma_1^2) + (1 - \alpha) N(\mu_2, \sigma_2^2)$$

and

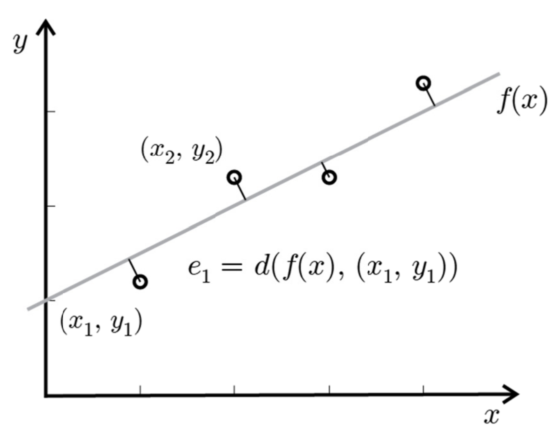
$$p_Y(y) = \beta N(\mu_1, \sigma_1^2) + (1 - \beta) N(\mu_2, \sigma_2^2)$$

where $\alpha \in (0,1)$, $\beta \in (0,1)$, $\mu_1 \in \mathbb{R}$, $\mu_2 \in \mathbb{R}$, $\sigma_1 \in \mathbb{R}^+$and $\sigma_2 \in \mathbb{R}^+$ are unknown parameters. $N(\mu, \sigma^2)$ is a univariate Gaussian distribution with mean μ and variance σ^2.

a) (5 points) Derive update rules of an EM algorithm for estimating , μ_1, μ_2, σ_1, and σ_2 based only on data set D_Y.
b) (15 points) Derive update rules of an EM algorithm for estimating α, β, μ_1, μ_2, σ_1, and σ_2 based on data set D.

Hint: In both cases the algorithm should follow the principle of maximizing the expected likelihood of complete data, as shown in class.

Problem 6 (20 points) Consider the problem of linear regression in which the objective function is to minimize the sum of squared distances to the fitting line, as shown in the figure below. In the figure, $d(f(x), (x_0, y_0))$ represents the Euclidean distance from point (x_0, y_0) to the line $f(x)$. Formulate the optimization problem and solve it as far as you can make it. Assume you are given a data set $D = \{(x_i, y_i)\}_{i=1}^n$, where $x \in \mathbb{R}$, and $y \in \mathbb{R}$.



Homework policies:

Your assignment must be typed; for example, in Latex, Microsoft Word, Lyx, etc. Images may be scanned and inserted into the document if it is too complicated to draw them properly. Submit a single pdf document or if you are attaching your code submit your code together with the typed (single) document as one .zip file.

All code (if applicable) should be turned in when you submit your assignment.

Policy for late submission assignments: Unless there are legitimate circumstances, late assignments will be accepted up to 5 days after the due date and graded using the following rule:

on time:	your score $\times$ 1
1 day late:	your score $\times$ 0.9
2 days late:	your score $\times$ 0.7
3 days late:	your score $\times$ 0.5
4 days late:	your score $\times$ 0.3
5 days late:	your score $\times$ 0.1

For example, this means that if you submit 3 days late and get 80 points for your answers, your total number of points will be $80 \times 0.5 = 40$ points.

All assignments are individual, except when collaboration is explicitly allowed. All the sources used for problem solution must be acknowledged, e.g. web sites, books, research papers, personal communication with people, etc. Academic honesty is taken seriously; for detailed information see Indiana University Code of Student Rights, Responsibilities, and Conduct.

Good luck!