ESE 524 - Homework 5

Problems
Assigned date: 03/27/19
Due date: 04/09/19
Total Points: 100 + 20 (extra credits)

1) **Bayesian Statistics**
The observed random variable is x. We want to estimate the parameter λ. The probability density of x as a function of λ is

$$p_{x|\lambda}(X|\lambda) = \begin{cases} \lambda e^{-\lambda X}, & X \geq 0, \lambda \geq 0, \\ 0, & X < 0. \end{cases}$$

The prior density of λ depends on two parameters: $n_0 \in \mathbb{Z}_{>0}$ and l_0 such that

$$p_{\lambda|n_0,l_0}(\lambda|n_0, l_0) = \begin{cases} \frac{l_0^{n_0}}{\Gamma(n_0)} e^{-\lambda l_0} \lambda^{n_0 - 1}, & \lambda \geq 0 \\ 0, & \lambda < 0. \end{cases}$$

a) **(10 pts)** Find $\mathbb{E}(\lambda)$ and $\mathrm{var}(\lambda)$ before any observations are made.
Hint: $\Gamma(n_0) = (n_0 - 1)!$

b) **(10 pts)** Assume that one observation is made. Find $p_{\lambda|x}(\lambda|X)$. Find the Bayesian MMSE estimate of λ, and denote it as $\hat{\lambda}_{\mathrm{MMSE}}$. Find $\mathbb{E}_x(\hat{\lambda}_{\mathrm{MMSE}} - \lambda)^2]$

c) **(10 pts)** Now assume that n independent observations are made. Denote these n observations by the vector $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^T$. Verify that

$$p_{\lambda|x}(\lambda|\boldsymbol{X}) = \begin{cases} \frac{l_* n_*}{\Gamma(n_*)} e^{-\lambda l_*} \lambda^{n_* - 1}, & \lambda \geq 0 \\ 0, & \lambda < 0. \end{cases}$$

where $l_* = l + l_0$ and $n_* = n + n_0$ with $l = \sum_{i=1}^{n} x_i$. Find $\hat{\lambda}_{\mathrm{MMSE}}$.

d) **(10 pts)** Also find the maximum a posterior estimate of λ assuming that there are n i.i.d observations as in part (c). Denote this estimator as $\hat{\lambda}_{\mathrm{MAP}}$. Is $\hat{\lambda}_{\mathrm{MAP}} = \hat{\lambda}_{\mathrm{MMSE}}$? Justify your answer.

2) **Lecture Notes, l4.pdf. Page #60**:
Define the loss function

$$L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\boldsymbol{x})) = \boldsymbol{c}^T \boldsymbol{a}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\boldsymbol{x}))$$

where $\boldsymbol{c} = [c_1, \ldots, c_p]^T$, $c_i > 0 \quad \forall i$ and a_i is the absolute error between i-th component of $\boldsymbol{\theta}$ and the i-th component of $\hat{\boldsymbol{\theta}}(\boldsymbol{x})$, i.e., $a_i(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\boldsymbol{x})) = |\theta_i - \hat{\theta}_i(\boldsymbol{x})|$.

(a) **(10 pts)** Show that the conditional risk may be written as

$$\int p(\boldsymbol{\theta}|\boldsymbol{x}) \boldsymbol{c}^T \boldsymbol{a}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\boldsymbol{x})) d\boldsymbol{\theta} = \sum_{i=1}^{p} c_i \beta$$

where

$$\beta = -\int_{-\infty}^{\hat{\theta}_i(\boldsymbol{x})} p(\theta_i|\boldsymbol{x})(\theta_i - \hat{\theta}_i(\boldsymbol{x})) d\theta_i + \int_{\hat{\theta}_i(\boldsymbol{x})}^{\infty} p(\theta_i|\boldsymbol{x})(\theta_i - \hat{\theta}_i(\boldsymbol{x})) d\theta_i.$$

(b) **(10 pts)** Show that the minimizing estimator $\hat{\theta}_i(\boldsymbol{x})$ is the median of the conditional density $f(\theta_i|\boldsymbol{x})$.

3) **Bayesian MSE**
(20 pts) In fitting a line through experimental data, we assume the model

$$x[n] = A + Bn + w[n], \quad \forall -M \le n \le M$$

where $w[n]$ is white Gaussian noise with variance σ^2. If we have some prior knowledge of the slope B and intercept A, such as

$$\begin{bmatrix} A \\ B \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} A_0 \\ B_0 \end{bmatrix}, \begin{bmatrix} \sigma_A^2 & 0 \\ 0 & \sigma_B^2 \end{bmatrix} \right),$$

find the MMSE estimator of A and B as well as the Bayesian MSE. Assume that A and B are independent of $w[n]$. Which parameter will benefit most from the prior knowledge?

4) **Kalman Filter (requires MATLAB)**
In Kalman filter, at every step, we update the estimation of the current state β_k using the estimate from the previous state β_{k-1} and only the observation from the current state y_k. Therefore, a Kalman filter is also a recursive filter. Now suppose that we have the following state and measurement equations:

$$\text{State Model}: \beta_k = \beta_{k-1}, \quad k = 1, 2, \ldots$$
$$\text{Measurement Model}: y_k = \beta_k + w_k, \quad k = 1, 2, \ldots$$

where β_k is a scalar state at time k and it does not vary, y_k is the observation at time k, and $w_k \sim \mathcal{N}(0, \sigma_{w_k}^2)$. Here $\{w_k,\ k = 1, 2, \ldots\}$ are mutually independent. The prior pdf for the initial state is $\beta_0 \sim \mathcal{N}(0, \sigma_{\beta_0}^2)$.

a) (**5 pts**) Write the Kalam filter equations for this linear system.

b) (**5 pts**) From another perspective, in this problem, the state is unchanged. It means that we actually keep estimating β_0 at every time step k, when we recevied the measurement y_k. Now, suppose that we are at time step k, i.e., we have observations $\boldsymbol{y}_{1:k} \overset{\text{def}}{=} \{y_1, y_2, \ldots, y_k\}$. Give the posterior distribution of β_0 (i.e., β_k).

Hint: Write it as a Gaussian linear model $\boldsymbol{y}_{1:k} = \boldsymbol{H}_k \beta_0 + \boldsymbol{W}_k$, and specify the $\boldsymbol{H}_k$ and $\boldsymbol{W}_k$, then you can use the formula from the lectures.

c) (**10 pts**) Now let $\sigma_{\beta_0}^2 = 100$, $\sigma_{w_k}^2 = 25$, and

$$\boldsymbol{y}_{1:10} = \{16.7549, 25.9058, 16.2077, 14.4519, 15.7722, 17.1367, 17.2066, 20.8919, 19.0157, 22.9322\}.$$

For time $k = 1, 2, 3, \ldots, 10$, use MATLAB to compute the MMSE estimates of β_k using the formulas in a) and b). Compare whether they are equal or not.

5) **Extra Credit**:
(**20 pts**) Come up with an example and solution illustrating one or more concepts from class so far. This example should be something you believe would be good to present in class to help other students understand a concept from the lectures. MATLAB (or other software) simulations are encouraged. Problems can be inspired by or explore applications from literature, but should not just copy the results of a paper. We will also accept larger "projects" which may take more time to complete but will be worth more points.