CS4070 – PART 2
EXERCISES RELATED TO LECTURE 3

(1) Assume $X_1, \dots, X_p$ are independent conditional on $\Theta_1, \dots, \Theta_p$. Suppose $X_i \mid \Theta_i = \theta_i \sim Unif(0, \theta_i)$. Consider estimation of the parameters $\Theta_1, \dots, \Theta_p$ based on data $X_1, \dots, X_p$.

(a) Model $\Theta_1, \dots, \Theta_p$ as independent with common density
$$f_\Theta(\theta) = \theta\lambda^2 e^{-\lambda\theta}\mathbf{1}_{[0,\infty)}(\theta)$$
for $\lambda > 0$. Show that $\Theta_1, \dots, \Theta_p$ are aposteriori independent. Find their posterior distribution and verify that
$$\mathbb{E}[\Theta_i \mid X_i] = X_i + 1/\lambda.$$

(b) If we use the posterior mean as estimator, then the performance of the estimator is highly depend on the choice of the hyperparameter λ. The method of empirical Bayes consists of plugging in an estimator for λ, based on
$$f_{X_1,\dots,X_p}(x_1, \dots, x_p) = \int f_{X_1,\dots,X_n \mid \Theta_1,\dots,\Theta_p}(x_p, \dots, x_p \mid \theta_1, \dots, \theta_p)$$
$$f_{\Theta_1,\dots,\Theta_p}(\theta_1, \dots, \theta_p)\,\mathrm{d}\theta_1, \dots \mathrm{d}\theta_p.$$
This is sometimes called the marginal likelihood, see also Section 3.9 in the book (which deals with a different statistical model, the underlying idea being the same as here).
Verify that $\mathbb{E}[X_i] = 1/\lambda$ and explain why $\hat{\Lambda} = 1/\bar{X}_n$ is an intuitively reasonable estimator for λ.
Hint: use $\mathbb{E}[X_i] = \mathbb{E}[\mathbb{E}[X_i \mid \Theta_i]]$.

(c) Determine an estimator for λ by what is called *marginal maximum likelihood* (sometimes also called maximum likelihood type II). This means we find λ as the/a maximiser of
$$\lambda \mapsto f_{X_1,\dots,X_p}(x_1, \dots, x_p).$$
Note that the dependence on λ is suppressed from the notation, but enters via the prior distribution.
To check your answer: you should find out that the marginal density of each X_i has the exponential distribution.

(d) Combine parts (a) and (b) (or (c)) to find empirical Bayes estimators for $\Theta_1, \dots, \Theta_p$. This means the estimator for λ is plugged in into the posterior mean found under part (a).

Date: November 23, 2023.

Solutions / hints to solutions

Note that I use Bayesian notation here, that is, instead of writing for example $f_X(x)$, I write $p(x)$.

(a). Note that the Θ_i follow a Gamma distribution $\Theta_i \sim Ga(2, \lambda)$. Further, it is easily seen that $\Theta_1, \ldots, \Theta_p$ are aposteriori independent and that the posterior distribution for Θ_i only depends on X_i. To find that distribution, we drop the index i from the notation and check that.

$$\begin{aligned} p(\theta \mid x) \propto p(x \mid \theta)p(\theta) &= \frac{1}{\theta}\mathbf{1}_{(0,\theta)}(x)\theta\lambda^2 e^{-\lambda\theta}\mathbf{1}_{(0,\infty)}(\theta) \\ &\propto \lambda^2 e^{-\lambda\theta}\mathbf{1}_{(x,\infty)}(\theta) \propto e^{-\lambda\theta}\mathbf{1}_{(x,\infty)}(\theta). \end{aligned}$$

Therefore,

$$p(\theta \mid x) = \frac{e^{-\lambda\theta}\mathbf{1}_{(x,\infty)}(\theta)}{\int_0^\infty e^{-\lambda\theta}\mathbf{1}_{(x,\infty)}(\theta)d\theta} = \lambda e^{-\lambda(\theta - x)}\mathbf{1}_{(x,\infty)}(\theta).$$

So each Θ_i is distributed as $X_i + Z_i$, where $\{Z_i,\, 1 \leq i \leq n\}$ is a sequence of IID $Exp(\lambda)$-distributed random variables, independent of all X_i. The posterior mean henceforth equals

$$\mathbb{E}[\Theta_i \mid X_i] = \frac{1}{\lambda} + X_i.$$

(b). Use the law of repeated expectation:

$$\mathbb{E}[X_i] = \mathbb{E}\mathbb{E}[X_i \mid \Theta_i] = \mathbb{E}[\Theta_i/2] = \frac{1}{2}\frac{2}{\lambda} = \frac{1}{\lambda}.$$

Now we can estimate λ by $\hat{\lambda} = 1/\bar{X}_n$.

(c.) Verify that $p(x_1, \ldots, x_n; \lambda) = \lambda^n e^{-\lambda\sum_{i=1}^p x_i}$, which is maximal for $\hat{\lambda} = 1/\bar{x}_n$. For the verification, note that

$$p(x_1, \ldots, x_n; \lambda) = \int \prod_{i=1}^p p(x_i \mid \theta_i)p(\theta_i; \lambda)\,\mathrm{d}x_i = \prod_{i=1}^p \int p(x_i \mid \theta_i)p(\theta_i; \lambda)\,\mathrm{d}x_i.$$

(d). Combining (a) and (b) gives $\hat{\theta}_i^{EB} = \bar{X}_n + X_i$ (just plug-in the emp. Bayes estimator for λ).