Statistics for Data Science — Exercise 12

Wintersemester 2023/24

Please complete these problems before the exercise session on Tuesday 23 January, 2024, 8:30. Please be prepared to present your solutions to any problems that you completed successfully.

1. Consider the following multiplicative noise model
$$y_j = a_j x_j, \quad 1 \leq j \leq n,$$
where $y, x, a \in \mathbb{R}^n$, and assume that a is a log-normally distributed multiplicative noise vector with independent components, that is, $\log a_j \sim \mathcal{N}(\log a_0, \sigma^2)$. Furthermore, a is assumed to be independent of x. By taking the logarithm, the noise model becomes additive. Using this observation, derive the likelihood density $f(y|x)$ for such $x \in \mathbb{R}^n$ that $x_j > 0$ for all $j = 1, \ldots, n$.

2. Consider the linear model
$$y = Ax + \eta,$$
where $x \in \mathbb{R}^d$ is the unknown, $y \in \mathbb{R}^k$ is the observation, $\eta \in \mathbb{R}^k$ is additive measurement noise, and $A \in \mathbb{R}^{k\times d}$ is the matrix modeling the measurement. Moreover, suppose that the noise distribution is given by $\eta \sim \mathcal{N}(0, \sigma^2 I)$ and the prior distribution by $x \sim \mathcal{N}(x_0, \gamma^2 I)$, where $\sigma > 0$, $\gamma > 0$, and $x_0 \in \mathbb{R}^d$.
   (a) Form the posterior density $f(x|y)$.
   (b) Notice that the *maximum a posteriori* (MAP) estimator is precisely the minimizer of $-\log(f(x|y))$. Using this observation, show that the MAP estimator $\hat{x}_{\mathrm{MAP}}$ is the solution to
$$(A^{\mathrm{T}}A + \lambda^2 I)\hat{x}_{\mathrm{MAP}} = A^{\mathrm{T}}y + \lambda^2 x_0,$$
   where $\lambda = \frac{\sigma}{\gamma}$.

3–4. (This task is worth 2 points.) Let $x, y, \eta \in \mathbb{R}$ and consider a simple linear model
$$y = \frac{1}{2}x + \eta$$
with additive noise $\eta \sim \mathcal{N}(0, 1)$. Assume that the prior model for the unknown is also Gaussian $x \sim \mathcal{N}(0, \frac{1}{\alpha})$, where $\alpha > 0$ is poorly known. It is possible to write the conditional prior for x, given α, as
$$f(x|\alpha) = \frac{\alpha^{1/2}}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\alpha x^2\right).$$
Since the parameter α is not known, it is part of the inference problem. Assume that we set the following hyperprior density for the parameter α:
$$f(\alpha) = \begin{cases} \sqrt{\frac{2}{\pi}} \exp(-\frac{1}{2}\alpha^2) & \text{if } \alpha > 0, \\ 0 & \text{if } \alpha \leq 0. \end{cases}$$

The exercises continue on the next page!

(a) Show that the posterior density for $(x, \alpha)|y$ is given by

$$f(x, \alpha|y) \propto \alpha^{1/2} \exp\left(-\frac{1}{2}\left(y - \frac{1}{2}x\right)^2 - \frac{1}{2}\alpha x^2 - \frac{1}{2}\alpha^2\right) \quad \text{for } x \in \mathbb{R},\ \alpha > 0,$$

where the implied coefficient does not depend on x or α.

(b) Show that $(x, \alpha) = (1, 1/2)$ is the *maximum a posteriori* (MAP) estimate when we observe $y = 3/2$.

You may assume in parts (a) and (b) that η and (x, α) are independent.

Hint: Task (b) can be solved by hand, but you may alternatively solve the optimization problem numerically.