

Return your written solutions either in person or by email
to veska.kaarnioja@fu-berlin.de by Monday 11 July, 2022, 12:15

1. Suppose $\rho_1 \sim \mathcal{N}(m_1, \sigma_1^2)$ and $\rho_2 \sim \mathcal{N}(m_2, \sigma_2^2)$. Solve the Kullback–Leibler divergence

$$d_{\text{KL}}(\rho_1 \parallel \rho_2) = \int_{\mathbb{R}} \log \left(\frac{\rho_1(x)}{\rho_2(x)} \right) \rho_1(x) dx.$$

2. Let $\rho \sim \mathcal{N}(0, 1)$ and assume that the posterior density is given by

$$\rho^y(x) = 0.6 \cdot \rho(x - 3) + 0.4 \cdot \rho(x + 2).$$

Using MATLAB try to find the KL-optimal approximation for ρ^y in the class of $\mathcal{B} = \{\mathcal{N}(m, 1) \mid m \in \mathbb{R}\}$, that is, find m that solves

- (a) $\inf_{m \in \mathbb{R}} d_{\text{KL}}(\mathcal{N}(m, 1) \parallel \rho^y)$
- (b) $\inf_{m \in \mathbb{R}} d_{\text{KL}}(\rho^y \parallel \mathcal{N}(m, 1))$

What difference do you find between the two approximations?

Hint: You are not required to do anything too sophisticated in this task. Any reasonable minimization procedure or quadrature rule is OK.

3. Show rigorously that $d_{\text{KL}}(\rho_1 \parallel \rho_2) = 0$ if and only if $\rho_1 = \rho_2$ a.e.