Final exam: Model Based Statistical Learning (1 hour)

December 4 2023

Recall that the p.d.f. of a Gaussian mixture model is

$$p_\theta(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k, \Sigma_k)$$

1. Detail the model parameters θ and give their constraints, if any.

2. Describe the interpretation of GMM as a latent variable model by introducing a latent variable.

3. How can you estimate the parameters ? Briefly describe the steps of the algorithm (without deriving all the computations, but write what you have to compute in each step).

4. Write the generative model of LDA and give the constraints of the parameters.

5. Compare GMM and LDA.

6. Consider now that we want to consider a mixture of two Bernoulli distributions. Write the probability distribution function of this mixture. Recall that a Bernoulli distribution is defined as

$$p(x|\theta) = \theta^x (1 - \theta)^{1-x}, \quad \text{with } x \in \{0, 1\}, \theta \in [0, 1]$$