Bayesian Methods for Machine Learning

Revisions

Exercise 1

Let

- $\mathcal{X} = \{x_n \in \mathbb{R}\}_{n=1}^N$ denote a set of observed random variables,
- $\mathcal{S}_1 = \{s_{1,n} \in \mathbb{R}\}_{n=1}^N$ and $\mathcal{S}_2 = \{s_{2,n} \in \mathbb{R}\}_{n=1}^N$ tow sets of latent random variables,
- $\theta = \{v_{j,n} \in \mathbb{R}_+^\star\}_{i=1,j=1}^{N,2}$ a set of known deterministic model parameters.

Question 1 We consider that for all $(j,n) \in \{1,2\} \times \{1,...,N\}$, the latent variables $s_{j,n}$ are mutually independent and follow a Gaussian prior distribution with zero mean and variance $v_{j,n}$. Given this information, write the factorization of the joint prior distribution $p(\mathcal{S}_1, \mathcal{S}_2; \theta)$.

We consider the following observation model:

$$x_n = s_{1,n} + s_{2,n}, \qquad \forall n \in \{1, ..., N\}. \tag{1}$$

Question 2 Give the expression of the following distributions:

- $p(x_n | s_{1,n}; \theta)$,
- $p(x_n | s_{2,n}; \theta)$,
- $p(x_n | s_{1,n}, s_{2,n})$.

Question 3 Using the plate notation, draw the Bayesian network corresponding to this model.

Question 4 Write the joint distribution of all the variables in this Bayesian network as a product of conditional distributions.

Question 5 Show that $p(\mathcal{X}; \theta) = \prod_{n=1}^{N} p(x_n; \theta)$.

Question 6 Show that $p(\mathcal{S}_1, \mathcal{S}_2 | \mathcal{X}; \theta) = \prod_{n=1}^{N} p(s_{1,n}, s_{2,n} | x_n; \theta)$.

Question 7 Show that $p(\mathcal{S}_j | \mathcal{X}; \theta) = \prod_{n=1}^{N} p(s_{j,n} | x_n; \theta)$ for $j \in \{1, 2\}$.

Question 8 Compute the posterior $p(s_{j,n} | x_n; \theta)$ for $j \in \{1, 2\}$.

Question 9 Propose a point estimate for the latent variable $s_{j,n}$, $j \in \{1, 2\}$.

Appendix

Gaussian distribution The probability density function (pdf) of the Gaussian distribution is given by

$$\mathcal{N}(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \tag{2}$$

where $x \in \mathbb{R}$ is the Gaussian random variable, $\mu = \mathbb{E}[x] \in \mathbb{R}$ is the mean and $\sigma^2 = \mathbb{E}[(x-\mu)^2] \in \mathbb{R}_+^*$ is the variance.