Advanced Probabilistic Machine Learning and Applications

Caterina De Bacco

April 20, 2021

1 Tutorial 1: Introduction to probabilistic ML

Exercise 1: Multivariate Gaussian

Given a data set $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}^\top$ in which the observations $\{\mathbf{x}_n\}$ are assumed to be drawn independently from a multivariate Gaussian distribution, i.e. $\mathbf{x}_1, \dots, \mathbf{x}_N \sim \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$:

1. Estimate the mean and covariance parameters $\boldsymbol{\mu}_x$ and $\boldsymbol{\Sigma}_x$, by maximum likelihood.

   Hints: (i) $|A^{-1}| = 1/|A|$, (ii) $\mathrm{Tr}[AB] = \mathrm{Tr}[BA]$, (iii) $\frac{\partial}{\partial A} \log |A| = A^{-\top}$, (iv) $\frac{\partial}{\partial A} Tr[AB] = B^\top$

2. Assume the covariance matrix $\boldsymbol{\Sigma}_x$ to be known and a Gaussian prior over the mean parameter $\boldsymbol{\mu}_x$ with mean $\boldsymbol{\mu}_0$ and identity covariance matrix, i.e. $\mathcal{N}(\boldsymbol{\mu}_x|\boldsymbol{\mu}_0, \mathbf{I})$. Compute the distribution a posteriori of the mean parameter μ_x given the observed data $\mathbf{X}$, i.e. $p(\boldsymbol{\mu}_x|\mathbf{X}, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_x)$, and its *Maximum a posteriori* (MAP) solution.

Exercise 2: Categorical distribution

Given a data set $\mathbf{X} = \{x_1, \dots, x_N\}^\top$ in which the observations $x_n \in \{1, \dots, k\}$ are assumed to be drawn independently from a Categorical distribution, i.e., $x_1, \dots, x_N \sim Categorical(x|\pi_1, \dots, \pi_k)$:

1. Estimate the parameters, i.e., the category probabilities $\{\pi_k\}$ by maximum likelihood.

   Hint: (i) Categorical distribution $p(x \mid \{\pi_k\}) = \prod_{k=1}^{K} \pi_k^{[x=k]}$, with $[x = k] = \begin{cases} 1 \text{ if } x = k \\ 0 \text{ otherwise} \end{cases}$

2. Assume a Dirichlet prior over the category probabilities $\{\pi_k\}$ with hyperparameter $\alpha = (\alpha_1, ..., \alpha_K)$, i.e., $\pi_1, \dots, \pi_k \sim Dirichlet(\pi_1, \dots, \pi_k|\alpha)$. Compute the distribution a posteriori of the category probabilities $\{\pi_k\}$ given the observed data $\mathbf{X}$, i.e., $p(\pi_1, \dots, \pi_k|\mathbf{X}, \alpha)$.

   Hint: (i) Dirichlet distribution: $p\left(\pi_1, \dots, \pi_K|\alpha_1, \dots, \alpha_K\right) \propto \prod_{k=1}^{K} \pi_k^{\alpha_k - 1}$