Advanced Probabilistic Machine Learning and Applications

Martina Contisciani and Caterina De Bacco

October 23, 2021

Tutorial 1: Introduction to probabilistic machine learning

Exercise 1: Multivariate Gaussian

Given a dataset $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}^\top$ in which the observations $\{\mathbf{x}_n\}$ are assumed to be drawn independently from a $K-$dimensional multivariate Gaussian distribution, i.e. $\mathbf{x}_n \sim \mathcal{N}_K(\mathbf{x}_n|\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$ $\forall n = 1, \ldots, N$:

1. Estimate the mean and covariance parameters $\boldsymbol{\mu}_x$ and $\boldsymbol{\Sigma}_x$, by *maximum likelihood* (ML).
2. Assume the covariance matrix $\boldsymbol{\Sigma}_x$ to be known and the existence of a multivariate Gaussian prior over the mean parameter $\boldsymbol{\mu}_x$ with mean $\boldsymbol{\mu}_0$ and identity covariance matrix, i.e. $\mathcal{N}_K(\boldsymbol{\mu}_x|\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ with $\boldsymbol{\Sigma}_0 = \mathbf{I}$. Compute the distribution a posteriori of the mean parameter $\boldsymbol{\mu}_x$ 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 dataset $\mathbf{X} = \{x_1, \ldots, x_N\}^\top$ in which the observations $x_n \in \{1, \ldots, K\}$ are assumed to be drawn independently from a Categorical distribution, i.e. $x_n \sim Categorical(x_n|\pi_1, \ldots, \pi_K)$ $\forall n = 1, \ldots, N$:

1. Estimate the parameters, i.e. the category probabilities $\{\pi_k\}$ by *maximum likelihood* (ML).
2. Assume a Dirichlet prior over the category probabilities $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_K)$ with hyperparameter $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_K)$, i.e. $\boldsymbol{\pi} \sim Dirichlet(\boldsymbol{\pi}|\boldsymbol{\alpha})$. Compute the distribution a posteriori of the category probabilities $\{\pi_k\}$ given the observed data $\mathbf{X}$, i.e. $p(\pi_1, \ldots, \pi_K|\mathbf{x}, \boldsymbol{\alpha})$.