Technische Universität Berlin
Fakultät IV – Elektrotechnik und Informatik

Probabilistic and Bayesian Modelling in Machine Learning and Artificial Intelligence

Manfred Opper, Théo Galy–Fajou

Summer Term 2018

Problem Sheet 2

Solutions to be discussed in the tutorial on Tuesday, May 22, 2018

Problem 1 – EM algorithm for a Poisson mixture model

Consider a mixture model for a integer valued random variable $n \in \{0, 1, 2, \ldots\}$ given by the distribution

$$P(n|\boldsymbol{\theta}) = \sum_{j=1}^{M} P(j)\ P(n|\theta_j) = \sum_{j=1}^{M} P(j)\ e^{-\theta_j} \frac{\theta_j^n}{n!}\,,$$

where the component probabilities $P(n|\theta_j)$ are Poisson distributions. Based on a data set of i.i.d. samples $D = (n_1, n_2, \ldots, n_N)$ we want to estimate the parameters $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_M, P(1), \ldots, P(M))$ of this mixture model.

(a) Derive an expression for the *Maximum Likelihood* estimate of θ_1 for $M = 1$, where all obervations come from the same Poisson distribution.

(b) For $M > 1$ the maximum likelihood estimates of the parameters are to be determined using an EM algorithm. Give explicit formulas for the update of θ_j and $P(j)$.

Hint: For the E-step (see the lecture), compute

$$\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\theta}_t) = -\sum_{i=1}^{N}\sum_{j=1}^{M} P_t(j|n_i) \ln\left(P(n_i|\theta_j)\, P(j)\right),$$

where $P_t(j|n_i)$ is the responsibility of component j for generating data point n_i, computed with the current values of the parameters. For the M-step, minimise $\mathcal{L}$ with respect to θ_j and $P(j)$.

Problem 2 – Bayesian estimation for the Poisson distribution

Consider again the Poisson distribution for an integer valued random variable $n \in \{0, 1, 2, \ldots\}$

$$P(n|\theta) = e^{-\theta}\frac{\theta^n}{n!} \,,$$

(a) Write the Poisson distribution in the *exponential family* form

$$P(n|\theta) = f(n) \exp\left[\psi(\theta)\phi(n) + g(\theta)\right]$$

(b) Use this exponential family representation to show that the *conjugate prior* for the Poisson distribution is given by the *Gamma density*

$$p(\theta|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}$$

where α, β are hyperparameters.

(c) Assume that we observe Poisson data $D = (n_1, n_2, \ldots, n_N)$. Write down the posterior distribution $p(\theta|D)$ assuming the *Gamma* prior. What are *posterior mean* and MAP estimators for θ ?

(d) Compute the *posterior variance* for large N and compare your result with the asymptotic frequentist error of the maximum likelihood estimator. **Hint:** For the computation of the frequentist error use the *Fisher Information* $J(\theta) \doteq E[(\frac{d \ln P(n|\theta)}{d\theta})^2]$ where the expectation is over the probability distribution $P(n|\theta)$.