N. Bertschinger
M. Kaschube
V. Ramesh

Machine Learning I
Exercise Sheet 4
Due on Mon, November 28, 12:15

If you turn in your solutions via email, please send to:
kamyshanska@fias.uni-frankfurt.de
Include [ML1-1617] in the subject line, turn in a single PDF file and include your name in the file name.

Exercise 1. *For the one-dimensional case, show that the product of two normal distributions with means μ_1, μ_2 and variances σ_1^2, σ_2^2 is proportional to a normal distribution with mean between the original two means and variance smaller than either of the original variances.*

2 points

Exercise 2. *Let $p(x|\mu)$ be a univariate Gaussian $\mathcal{N}(\mu, \sigma^2)$ with unknown parameter mean, which is also assumed to follow a Gaussian $\mathcal{N}(\mu_0, \sigma_0^2)$. From the theory exposed before we have*

$$p(\mu|X) = \frac{p(X|\mu)p(\mu)}{p(X)} = \frac{1}{\alpha}\prod_{k=1}^{N} p(x_k|\mu)p(\mu)$$

where for a given training data set, X, $p(X)$ is a constant denoted as α. Write down the explicit expression for $p(\mu|X)$.

1 points

Exercise 3. *Show that, given a number of samples, N, the posterior $p(\mu|X)$ turns out to be also Gaussian, that is*

$$p(\mu|X) = \frac{1}{\sigma_N\sqrt{2\pi}} \exp\left(-\frac{(\mu-\mu_N)^2}{2\sigma_N^2}\right)$$

with mean value

$$\mu_N = \frac{N\sigma_0^2\bar{x}_N + \sigma^2\mu_0}{N\sigma_0^2 + \sigma^2}$$

and variance

$$\sigma_N^2 = \frac{\sigma^2\sigma_0^2}{N\sigma_0^2 + \sigma^2}$$

where $\bar{x}_N = \frac{1}{N}\sum_{k=1}^{N} x_k$. In the limit of large N, what happens to the mean value μ_N and to the standard deviation σ_N?

4 points

N. Bertschinger
M. Kaschube
V. Ramesh

Machine Learning I
Exercise Sheet 4
Due on Mon, November 28, 12:15

Exercise 4. *Plot the posterior distribution* $p(\mu|X)$ *from Exercise 3 in one graph for various* N*. The largest* N *should be at least as large as* $N = 100$*. To compute* $\bar{x}_N$ *generate data* $X = \{x_1, \ldots, x_N\}$ *using a pseudorandom number generator following a Gaussian pdf with mean value equal* $\mu = 2$ *and variance* $\sigma^2 = 4$*. The mean value is assumed to be unknown and the prior pdf is also a Gaussian with* $\mu_0 = 0$ *and* $\sigma_0^2 = 8$*. Also include the prior in this plot and describe what happens when increasing* N*.*

2 points

Exercise 5. *Show that the posterior pdf estimate in the Bayesian inference task, for independent variables, can be computed recursively, that is,*

$$p(\theta|x_1, \ldots, x_N) = \frac{p(x_N|\theta)p(\theta|x_1, \ldots, x_{N-1})}{p(x_N|x_1, \ldots, x_{N-1})} \ .$$

You may either do this for the general expression above or for the example of Gaussian distributions.

1 points