Statistical Machine Learning: Asssignment 3

Joris van Vugt, s4279859

November 24, 2016

Exercise 1 – Bayesian linear regression

1. The predictive distrubtion after observing these two data points is

$$p(t|x, \boldsymbol{x}, \boldsymbol{t}) = \mathcal{N}(t|m(x), s^2(x))$$

. Using

$$\begin{aligned}
\boldsymbol{S}_N^{-1} &= \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix} + N\beta \begin{pmatrix} 1 & \bar{\mu}_x \\ \bar{\mu}_x & \bar{\mu}_{xx} \end{pmatrix} \\
&= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} + 20 \begin{pmatrix} 1 & 0.5 \\ 0.5 & 0.26 \end{pmatrix} \\
&= \begin{pmatrix} 22 & 10 \\ 10 & 7.2 \end{pmatrix}
\end{aligned}$$

we can compute the mean and standard deviation of the predictive distribution for each x

$$\begin{aligned}
m(x) &= N\beta \begin{pmatrix} 1 & x \end{pmatrix} \boldsymbol{S}_N \begin{pmatrix} \bar{\mu}_t \\ \bar{\mu}_{xt} \end{pmatrix} \\
&= 2 \times 10 \begin{pmatrix} 1 & x \end{pmatrix} \begin{pmatrix} 22 & 10 \\ 10 & 7.2 \end{pmatrix} \begin{pmatrix} 0.15 \\ -0.95 \end{pmatrix} \\
&= 20 \begin{pmatrix} 1 & x \end{pmatrix} \begin{pmatrix} 22 & 10 \\ 10 & 7.2 \end{pmatrix} \begin{pmatrix} 0.15 \\ -0.95 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
s^2(x) &= \beta^{-1} + \begin{pmatrix} 1 & x \end{pmatrix} \boldsymbol{S}_N \begin{pmatrix} 1 \\ x \end{pmatrix} \\
&= 10^{-1} + \begin{pmatrix} 1 & x \end{pmatrix} \begin{pmatrix} 22 & 10 \\ 10 & 7.2 \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix}
\end{aligned}$$