Machine Learning
Exercise 8

Marc Toussaint
Machine Learning & Robotics lab, U Stuttgart
Universittsstrae 38, 70569 Stuttgart, Germany

June 26, 2013

1 Gaussian Processes

Consider a Gaussian Process prior $P(f)$ over functions defined by the mean function $\mu(x) = 0$, the γ-exponential covariance function

$$k(x, x') = \exp\{-|(x - x')/l|^\gamma\}$$

and an observation noise $\sigma = 0.1$. We assume $x \in \mathbb{R}$ is 1-dimensional. First consider the standard squared exponential kernel with $\gamma = 2$ and $l = 0.2$.

a) Write a routine to draw 10 random functions from the prior $P(f)$. For this, discretize $x \in [-1, 1]$ to 100 points, compute the covariance matrix for these 100 points and sample.

b) Assume we have two data points $(-0.5, 0.3)$ and $(0.5, -0.1)$. Display the posterior $P(f|D)$. For this, compute the mean posterior function $\widehat{f}(x)$ and the standard deviation function $\widehat{\sigma}(x)$ (on the 100 grid points) exactly as on slide 07:9, using $\lambda = \sigma^2$. Then plot $\widehat{f}$, $\widehat{f} + \widehat{\sigma}$ and $\widehat{f} - \widehat{s}$ to display the posterior mean and standard deviation.

c) Now display the posterior $P(y^*|x^*, D)$. This is only a tiny difference from the above (see slide 07:7). The mean is the same, but the variance of y^* includes additionally the obseration noise σ^2.

d) Sample 10 functions from the posterior $P(f|D)$ and display them.

e) Repeat a-d) for a kernel with $\gamma = 1$.

f) Optional: In case you have the code: compare the posterior $P(f|D)$ with standard Ridge regression using radial basis function features of width 0.2.