

# Machine Learning

## Exercise 11

Marc Toussaint

Machine Learning & Robotics lab, U Stuttgart  
Universitätsstraße 38, 70569 Stuttgart, Germany

July 1, 2016

### 1 Gaussian Processes

Consider a Gaussian Process prior  $P(f)$  over functions defined by the mean function  $\mu(x) = 0$ , the  $\gamma$ -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) 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  $\hat{f}(x)$  and the standard deviation function  $\hat{\sigma}(x)$  (on the 100 grid points) exactly as on slide 06:10, using  $\lambda = \sigma^2$ . Then plot  $\hat{f}$ ,  $\hat{f} + \hat{\sigma}$  and  $\hat{f} - \hat{\sigma}$  to display the posterior mean and standard deviation.

b) Now display the posterior  $P(y^*|x^*, D)$ . This is only a tiny difference from the above (see slide 06:8). The mean is the same, but the variance of  $y^*$  includes additionally the observation noise  $\sigma^2$ .

c) Repeat a) & b) for a kernel with  $\gamma = 1$ .