

**Machine Learning for Computer Vision**  
**Summer term 2017**

22. Mai 2017

Topic: Regression, Kernels and Gaussian Processes

**Exercise 1: Bayesian Update**

Consider a linear regression model with basis functions  $\phi(x)$  as presented in the lecture. We assume a Gaussian prior distribution for the weights:

$$p(\mathbf{w}) = \mathcal{N}(\mathbf{w}|m_0, S_0)$$

Suppose we have already observed  $N$  data points, so the posterior distribution is

$$p(\mathbf{w}|\mathbf{t}) = \mathcal{N}(\mathbf{w}|m_N, S_N)$$

with

$$m_N = S_N(S_0^{-1}m_0 + \sigma^{-2}\Phi^T\mathbf{t}) \quad \text{and} \quad S_N = S_0^{-1} + \sigma^{-2}\Phi^T\Phi.$$

Now, we observe a new data point  $(x_{N+1}, t_{N+1})$ . What is the new posterior?

**Exercise 2: Constructing kernels**

Let  $k_1$  and  $k_2$  be kernels,  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  an arbitrary function. Show that we can construct new kernels via

- a)  $k(x_1, x_2) = k_1(x_1, x_2) + k_2(x_1, x_2)$
- b)  $k(x_1, x_2) = k_1(x_1, x_2)k_2(x_1, x_2)$
- c)  $k(x_1, x_2) = f(x_1)k_1(x_1, x_2)f(x_2)$
- d)  $k(x_1, x_2) = \exp(k_1(x_1, x_2))$
- e)  $k(x_1, x_2) = x_1^T A x_2$ , where  $A$  symmetric, positive semi-definite  $n \times n$  matrix

**Exercise 3: Polynomial kernel**

Let  $x_i, x_j \in \mathbb{R}^2$

- a) Show (by induction) that  $k_d(x_i, x_j) = (x_i^T x_j)^d$  is a kernel for every  $d \geq 1$ .
- b) Find  $\phi_d(x)$  such that  $k_d(x_i, x_j) = \phi_d(x_i)^T \phi_d(x_j)$ .

c) Find  $\tilde{\phi}_2(x)$  for  $\tilde{k}_2(x_i, x_j) = (x_i^T x_j + d)^2$  ( $d > 0$ ).

#### Exercise 4: Gaussian Processes Regression (Programming)

Download the file *train\_data.txt*. The file contains pairs of inputs and targets you can use to train a Gaussian Process model. Can you fit a function to these points? Start by using a squared-exponential kernel with length scale  $l = 3$ ,  $\sigma_f = 1$  and  $\sigma_n = 0.5$ . Try different values for the length scale hyperparameter. How does the model change?