Fachgebiet Maschinelles Lernen
Institut für Softwaretechnik und theoretische Informatik
Fakultät IV, Technische Universität Berlin
Prof. Dr. Klaus-Robert Müller
Email: klaus-robert.mueller@tu-berlin.de

Exercise Sheet 5

A kernel function $k: \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ must satisfy the *Mercer's condition*, which verifies that for any sequence of data points $x_1, \ldots, x_n \in \mathbb{R}^d$ and coefficients $c_1, \ldots, c_n \in \mathbb{R}$ the inequality

$$\sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j k(x_i, x_j) \ge 0$$

is satisfied. If it is the case, the kernel is called a Mercer kernel.

Conversely, the representer theorem states that if k is a Mercer kernel on \mathbb{R}^d , then there exists a Hilbert space (i.e., a finite or infinite dimensional \mathbb{R} -vector space with norm and scalar product) \mathcal{F} , the so-called feature space, and a continuous map $\varphi : \mathbb{R}^d \to \mathcal{F}$, such that

$$k(x, x') = \langle \varphi(x), \varphi(x') \rangle_{\mathcal{F}}$$
 for all $x, x' \in \mathbb{R}^d$.

Exercise 1: Mercer Kernels $(3 \times 20 \text{ P})$

- (a) Show that the following are Mercer kernels.
 - i. $k(x, x') = \langle x, x' \rangle$
- ii. $k(x,x') = f(x) \cdot f(x')$ where $f: \mathbb{R}^d \to \mathbb{R}$ is an arbitrary continuous function
- (b) Let k_1, k_2 be two Mercer kernels, for which we assume the existence of a finite-dimensional feature map associated to them. Show that the following are again Mercer kernels.
 - i. $k(x, x') = k_1(x, x') + k_2(x, x')$
- ii. $k(x, x') = k_1(x, x') \cdot k_2(x, x')$
- (c) Show using the results above that the polynomial kernel of degree d, where $k(x, x') = (\langle x, x' \rangle + \vartheta)^d$ and $\vartheta \in \mathbb{R}^+$, is a Mercer kernel.

Exercise 2: The Feature Map $(4 \times 10 P)$

Consider the homogenous polynomial kernel k of degree 2 which is $k: \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, where

$$k(x,y) = \langle x, y \rangle^2 = \Big(\sum_{i=1}^2 x_i y_i\Big)^2.$$

- (a) Show that $\mathcal{F} = \mathbb{R}^3$ and $\varphi \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1^2 \\ \sqrt{2} x_1 x_2 \\ x_2^2 \end{pmatrix}$ are possible choices for feature space and feature map.
- (b) Consider the unit circle $C = \left\{ \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} ; \ 0 \le \theta < 2\pi \right\}$. Show that the image $\varphi(C)$ lies on a plane H in \mathbb{R}^3 .
- (c) Consider the plane $A = \left\{ \begin{pmatrix} t \\ s \end{pmatrix} ; t, s \in \mathbb{R} \right\}$. Find a point P in \mathcal{F} which is not contained in $\varphi(A)$.
- (d) Find a feature map associated to the kernel $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ with $k(x,y) = \langle x,y \rangle^2 = \Big(\sum_{i=1}^d x_i y_i\Big)^2$.