*Themen des statistischen maschinellen Lernens*
Wintersemester 2016/17
Humboldt-Universität zu Berlin

Dr. Martin Wahl

## Sheet 3

1. (a) Show that if $k_1, k_2 : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ are two positive definite kernels (see Exercise 4, Sheet 2 for the definition) and $a > 0$, then $ak_1$, $k_1 + k_2$ and $k_1 \cdot k_2$ are also positive definite kernels.

   (b) Show that if $k_n : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, $n \geq 1$, is a sequence of positive definite kernels which converge pointwise to a function $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$, then $k$ is also a positive definite kernel.

   (c) Let $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be a positive definite kernel. Show that $k' : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ defined by $k'(x, y) = k(x, y)/\sqrt{k(x, x)k(y, y)}$, if the denominator is non-zero, and $k'(x, y) = 0$, otherwise, is also a positive definite kernel.

   (d) Deduce that the map $k : \mathbb{R}^q \times \mathbb{R}^q \to \mathbb{R}, (x, y) \mapsto \exp(-\|x - y\|^2/(2\sigma^2))$ is a positive definite kernel, where $q \geq 1$ and $\sigma^2 > 0$.

2. Suppose that $X, X_1, \ldots, X_n$ are i.i.d. random variables taking values in $\mathbb{R}^p$ such that $\mathbb{E}[X] = 0$ and $\mathbb{E}[\|X\|^4] < \infty$. Let $\Sigma = \mathbb{E}[XX^T]$ and $\hat{\Sigma} = (1/n) \sum_{i=1}^n X_i X_i^T$. Show that

$$\mathbb{E}[\|\hat{\Sigma} - \Sigma\|_{HS}^2] = \frac{1}{n}\big(\mathbb{E}[\|X\|^4] - \|\Sigma\|_{HS}^2\big).$$

   In the case that each $X_i$ is Gaussian, show that

$$\mathbb{E}[\|X\|^4] - \|\Sigma\|_{HS}^2 = (\operatorname{tr}(\Sigma))^2 + \operatorname{tr}(\Sigma^2).$$

3. Let $0 \leq \theta_1 \leq \cdots \leq \theta_d \leq \pi/2$ be the principal angles between two $d$-dimensional subspaces $V$ and $W$ of a Hilbert space $(H, \langle \cdot, \cdot \rangle)$, and let $P_V$ and $P_W$ be the orthogonal projections onto $V$ and $W$ respectively.

   (a) Show that

$$\cos(\theta_1) = \max_{v \in V, w \in W} \frac{|\langle v, w \rangle|}{\|v\|\|w\|}.$$

   (*Hint*: Use that the first singular value $s_1(A)$ of a $d \times d$-matrix $A$ satisfies $s_1(A) = \max_{\|x\|=1} \|Ax\| = \max_{\|x\|,\|y\|=1} y^T Ax$.) State and prove a similar formula for $\cos(\theta_j)$, $j = 2, \ldots, d$.

   (b) Show that

$$\|P_V - P_W\|_{HS}^2 = 2 \sum_{j=1}^d \sin^2(\theta_j).$$