Machine Learning Theory
Tübingen University, WS 2016/2017
Assignment 5

Due: Monday, January 30th, 12:15

Hand in at beginning of tutorial!

The solutions to this assignment will be discussed in the tutorial on Monday, Jan 30. If you want your assignment to be marked, it has to be handed in at the beginning of the tutorial. If you can not make it to the tutorial yourself, please arrange for one of your classmates to hand in your assignment. In this assignment you can get 20 points. You can gain up to 20 bonus points by solving the last two problems.

To better understand the questions, refer to Lecture 11. To gain full marks, explain your answers carefully and prove your claims.

1. **Polynomial kernel of a second degree** In lecture 11 we claimed that $\mathcal{H}_k$ corresponding to the polynomial kernel of a second degree $k(x,y) = (\langle x,y\rangle_{\mathbb{R}^2} + 1)^2$, where $x, y \in \mathcal{X} = \mathbb{R}^2$, contains all the functions of the form

$$f(x) = v_1 x_1^2 + v_2 x_2^2 + v_3 x_1 x_2 + v_4 x_1 + v_5 x_2 + v_6, \quad x \in \mathcal{X}, v \in \mathbb{R}^6,$$

   i.e. all the polynomials up to 2-nd degree. The remaining bit to show this was to prove the following statement, which you are asked to prove now.

   Show that a linear span (a set of all the linear combinations of the vectors from a given set) of

$$\{(w_1^2, w_2^2, 2w_1w_2, 2w_1, 2w_2, 1) \colon w_1, w_2 \in \mathbb{R}\} \subset \mathbb{R}^6$$

   gives the whole $\mathbb{R}^6$. **(5 marks)**

2. **Reproducing kernels vs. feature maps.** Consider any space $\mathcal{X}$ and take any function $k \colon \mathcal{X} \times \mathcal{X} \to \mathbb{R}$. Show that, in order to prove that k is a kernel, it is enough to find (a) *any* vector space $\mathcal{H}$ with an inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}}$ and (b) a mapping $\psi \colon \mathcal{X} \to \mathcal{H}$, such that

$$k(x,y) = \langle \psi(x), \psi(y) \rangle_{\mathcal{H}}$$

   holds for all $x, y \in \mathcal{X}$. **(5 marks)**

3. **Gaussian kernel is indeed a valid kernel.** Consider $\mathcal{X} = \mathbb{R}$ and define *the Gaussian kernel* $k(x,y) = e^{-(x-y)^2/(2\sigma^2)}$ for any constant $\sigma^2 > 0$. Prove that this function is indeed a kernel. Hint: use Taylor's expansion for $\exp(x)$ function, which says that for all $x \in \mathbb{R}$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots.$$

   (10 marks)

4. **Gaussian kernel has an infinite dimensional RKHS.** Prove that RKHS $\mathcal{H}_k$ of a Gaussian kernel is infinite-dimensional, using following two facts.

   Fact 1. Consider any inner product space $\mathcal{H}$. Then a finite set of vectors $v_1, \dots, v_n \in \mathcal{H}$ is linearly independent *if and only if* the corresponding *Gram matrix* $G_V \in \mathbb{R}^{n \times n}$ with (i, j)-th element being $\langle v_i, v_j \rangle_{\mathcal{H}}$ has a non-zero determinant (i.e. is of a full rank).

   Fact 2. Take any $X_1, \dots, X_n \in \mathbb{R}$. Show that the matrix $G \in \mathbb{R}^{n \times n}$ with (i, j)-th element defined by $e^{-(X_i - X_j)^2/(2\sigma^2)}$ is of a full rank. **(10 marks)**

5. **Gram matrix of a Gaussian kernel is of a full rank.** Prove Fact 2 from the previous problem. **(20 marks)**