Wissenschaftliches Rechnen II/Scientific Computing II

Sommersemester 2016
Prof. Dr. Jochen Garcke
Dipl.-Math. Sebastian Mayer

universitätbonn

Exercise sheet 8 To be handed in on **Thursday, 16.06.2016**

Conditionally Positive Semi-Definite Kernels

The main idea of this exercise sheet is to give you a thorough understanding how conditional positive semi-definiteness (cpsd) is naturally related to regularized least-squares problems with penalty functionals J that have non-trivial kernel $\ker(J) = \{f : J(f) = 0\}$.

You will mostly consider the univariate cubic smoothing spline on the Sobolev space $W^2 := W^2([0,1])$ as introduced in Sheet 5, H2. We stress that in this setting cpsd is not needed from a practical point of view since we have nice psd reproducing kernels. Nevertheless this setting serves as a good demonstration and allows to avoid to much technicalities.

The most important exercise in this regard is H3, which we strongly recommend to work on. The group exercises should give you a profound preparation for H3.

Recall that for given data $(x_1, y_1), \ldots, (x_N, y_N)$ with pairwise different sample points $x_i \in [0,1]$, the cubic smoothing spline is given by $\hat{f} = \operatorname{argmin}_{f \in W^2} R_{\ell_2,\mathrm{reg}}(f)$. The regularized empirical risk is given by

$$R_{\ell_2,\mathrm{reg}}(f) = \frac{1}{N}\sum_{i=1}^{N}(y_i - f(x_i))^2 + \lambda J_2(f)$$

with penalty functional $J_2(f) = \|P_1 f\|^2_{W^2}$, where P_1 denotes the orthogonal projection onto $W_0^2 = \{f \in W^2 : f(0) = f'(0) = 0\}$.

1 Group exercises

G 1. Provide the proof details for Lemma 51 given in the lecture.

G 2. Let $R(x,y)$ and $R_1(x,y)$ be the kernels defined in Sheet 5, H2. Imitating the approach given in the lecture notes on p. 35/36, derive the system of linear equations which determines the solution $\hat{f}$. Then show that you can replace R_1 by R in the kernel matrix which appears in the derived linear system. **Hint:** Choose as the first two elements of your ONB the functions $\phi_1(x) = 1$, $\phi_2(x) = x$.

G 3. Choose two arbitrary distinct points t_1, t_2 from the set of sample points $\{x_1, \ldots, x_n\}$, say w.l.o.g. $t_1 = x_1$, $t_2 = x_2$. Let p_1, p_2 be the unique polynomials of degree 1 which solve $p_i(t_j) = \delta_{ij}$ for $i, j \in \{1, 2\}$, where δ_{ij} denotes the Kronecker delta (p_1, p_2 form a so-called *Lagrange basis* of $\Pi_1 = \operatorname{span}\{\phi_1, \phi_2\}$). As you have proved in G2 the smoothing spline has the form

$$\hat{f}(x) = \underbrace{\alpha_0 + \alpha_1 x}_{\text{affine part}} + \underbrace{\sum_{j=1}^{N} z_j R_1(x, x_j)}_{\text{kernel part}} \tag{8}$$

a) Show that the conditions $\sum_{i=1}^{N} z_i = \sum_{i=1}^{N} z_i x_i = 0$, which you have derived in G2, are equivalent to $\sum_{i=1}^{N} z_i p_1(x_i) = \sum_{i=1}^{N} z_i p_2(x_i) = 0$.

b) Use a) to show that the kernel part is contained in the set $P_1(V_0)$, where

$$V_0 := \{f \in W^2 : f(t_1) = f(t_2) = 0\}.$$

2 Homework

H 1. Let $\beta \in \mathbb{N}$. Show that the *thin plate spline* (TPS) kernel

$$k_\beta : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}, (x, y) \mapsto (-1)^{\beta+1} \|x - y\|_2^{2\beta} \log(\|x - y\|_2)$$

is conditionally positive semi-definite of order $\beta + 1$.

(2 Punkte)

H 2. The so-called *sigmoid kernel*

$$k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}, \quad (x, y) \mapsto k(x, y) = \tanh(\kappa x^T y + \nu),$$

where $\kappa >, \nu < 0$ and tanh denotes the tangens hyperbolicus, is often used in the context of artifical neural networks. Show that k is not conditionally positive semi-definite. **Hint:** Construct a counter example which exploits the fact that $\lim_{t \to \pm\infty} \tanh'(t) = 0$.

(4 Punkte)

H 3. (A different perspektive on cpsd)

Consider again the Sobolev space W_2. The starting point of $G2$ was to use the orthogonal decomposition

$$W^2 = \Pi_1 \oplus W_0^2,$$

which is reflected in the representation (8) of the smoothing spline $\hat{f}$. An alternative representation of $\hat{f}$ is of the form

$$\hat{f} = I_{(t_1,t_2)}(\hat{f}) + g,$$

where $I_{(t_1,t_2)}(\hat{f})(x) = p_1(x)\hat{f}(t_1) + p_2(x)\hat{f}(t_1)$ is the linear interpolation of $\hat{f}$ in the points t_1, t_2, and $g \in V_0$.

a) Using representation (8), determine $I_{(t_1,t_2)}(\hat{f})$ and g.

b) Determine the reproducing kernels of the subspaces V_0 and $P_1(V_0)$.

(10 Punkte)