Topics in Applied Mathematics: Infinitely Large Neural Networks, 3341.751
E. Ryu
Spring 2022

Homework 4
Due 5pm, Tuesday, April 26, 2022

Problem 1: *Non-strict representer theorem.* Let $\mathcal{X}$ be a nonempty set, $K\colon \mathcal{X}\times\mathcal{X}\to\mathbb{R}$ a PDK, $\mathcal{H}$ the corresponding RKHS, $X_1,\dots,X_N\in\mathcal{X}$, and $Y_1,\dots,Y_N\in\mathbb{R}$. Consider the optimization problem

$$\underset{f\in\mathcal{H}}{\text{minimize}}\quad L(\{(X_i,Y_i,f(X_i))\}_{i=1}^N)+Q(\|f\|_{\mathcal{H}})$$

where $Q\colon\mathbb{R}_+\to\mathbb{R}$ is a non-decreasing function. Show that, if a minimizer exists, there is a minimizer in

$$\mathrm{span}(\{K(X_i,\cdot)\}_{i=1}^N).$$

Problem 2: *All solutions of kernel ridge regression.* Let $\mathcal{X}$ be a nonempty set. Let $X_1,\dots,X_N\in\mathcal{X}$, $Y_1,\dots,Y_N\in\mathbb{R}$, $\lambda>0$, $K\colon\mathcal{X}\times\mathcal{X}\to\mathbb{R}$ be a PDK, $\mathcal{H}$ the corresponding RKHS, and $G\in\mathbb{R}^{N\times N}$ the kernel matrix defined as $G_{ij}=K(X_i,X_j)$ for $i,j\in\{1,\dots,N\}$. Consider

$$\underset{f\in\mathcal{H}}{\text{minimize}}\quad \frac{1}{N}\sum_{i=1}^N(f(X_i)-Y_i)^2+\lambda\|f\|_{\mathcal{H}}^2.$$

Let

$$\varphi^\star=(G+\lambda NI)^{-1}Y.$$

Show that (i)

$$f_v^\star(\cdot)=\sum_{j=1}^N(\varphi_j^\star+v_j)K(\cdot,X_j).$$

for $v\in\mathcal{N}(G)$, is the set of all solutions and that (ii) $f_u^\star=f_v^\star$ for all $u,v\in\mathcal{N}(G)$.

Hint. For (ii), show that $\|f_u^\star-f_v^\star\|_{\mathcal{H}}=0$.

Problem 3: Let $K\colon\mathcal{X}\times\mathcal{X}\to\mathbb{R}^{d\times d}$ and $k\colon(\mathcal{X}\times\{1,\dots,d\})\times(\mathcal{X}\times\{1,\dots,d\})\to\mathbb{R}$ such that

$$(K(x,x'))_{ij}=k((x,i),(x',j)).$$

Show that K is a mvPDK if and only if k is a (scalar-valued) PDK.

Problem 4: Let $\mathcal{X}$ be a nonempty set. Let $K_1\colon\mathcal{X}\times\mathcal{X}\to\mathbb{R}^{d\times d}$ and $K_2\colon\mathcal{X}\times\mathcal{X}\to\mathbb{R}^{d\times d}$ be mvPDKs. Define $K_3\colon\mathcal{X}\times\mathcal{X}\to\mathbb{R}^{d\times d}$ as

$$K_3(x,x')=K_1(x,x')\odot K_2(x,x'),\qquad\forall\,x,x'\in\mathcal{X},$$

where $\odot$ denotes the Hadamard product. Show that K_3 is an mvPDK.