STAT 672

Statistical Learning II

John Karasev
Homework 2
January 31, 2020

1 Kernel centering

Let $\mathcal{X}$ be a non empty set, K a kernel over $\mathcal{X}$ and H the RKHS with kernel K. Let $x_1, \ldots, x_n$, be n points in $\mathcal{X}$. Let K_c like "K centered" be another kernel over $\mathcal{X}$ defined by

$$K_c(x, y) = \langle K(., x) - \bar{f}(.), K(., y) - \bar{f}(.)\rangle_H, \text{ with } \bar{f}(.) = \frac{1}{n}\sum_{i=1}^{n} K(., x_i) \tag{1}$$

1. Verify that K_c is a positive definite kernel;
   Let H_c be the RKHS with reproducing kernel K_c.

   Proof.

   for any n in $\mathbb{N}$, $(a_1, ..., a_n)$ in $\mathbb{R}^n$,
   $(x_1, ..., x_n) \in \mathbb{X}^n$

   $$\begin{aligned}
   &\sum_{i,j=1}^{n} a_i a_j K_c(x_i, x_j) \\
   &= \sum_{i,j=1}^{n} a_i a_j \langle K(., x_i) - \bar{f}(.), K(., x_j) - \bar{f}(.)\rangle_{\mathcal{H}} \\
   &= \left\langle \sum_{i=1}^{n} a_i[K(., x_i) - \bar{f}(.)], \sum_{j=1}^{n} a_j[K(., x_j) - \bar{f}(.)] \right\rangle_{\mathcal{H}} \\
   &= \left\| \sum_{i=1}^{n} a_i[K(., x_i) - \bar{f}(.)] \right\|_{\mathcal{H}} \geq 0
   \end{aligned}$$

   □

2. (**) Verify that for any $f \in H_c$,

   $$\frac{1}{n}\sum_{l=1}^{n} f(x_l) = 0 \tag{2}$$

   $$f(x) = \sum_{m=1}^{n} \alpha_m K_c(x, y_m)$$

Proof.

$$\frac{1}{n}\sum_{l=1}^{n} f(x_l) = 0$$

$$\frac{1}{n}\sum_{l=1}^{n}\sum_{m=1}^{n} \alpha_m K_c(x_l, y_m) = 0$$

$$\frac{1}{n}\sum_{l=1}^{n}\sum_{m=1}^{n} \alpha_m \left\langle K(., x_l) - \frac{1}{n}\sum_{i=1}^{n} K(., x_i), K(., y_m) - \frac{1}{n}\sum_{i=1}^{n} K(., x_i) \right\rangle_{\mathcal{H}} = 0$$

$$\frac{1}{n}\left\langle \sum_{l=1}^{n} K(., x_l) - \sum_{l=1}^{n}\frac{1}{n}\sum_{i=1}^{n} K(., x_i), \sum_{m=1}^{n} \alpha_m K(., y_m) - \sum_{m=1}^{n}\alpha_m \frac{1}{n}\sum_{i=1}^{n} K(., x_i) \right\rangle_{\mathcal{H}} = 0$$

$$\left\langle \sum_{l=1}^{n} K(., x_l) - n\frac{1}{n}\sum_{i=1}^{n} K(., x_i), \sum_{m=1}^{n} \frac{\alpha_m K(., y_m)}{n} - \sum_{m=1}^{n}\frac{\alpha_m}{n} \sum_{i=1}^{n} \frac{K(., x_i)}{n} \right\rangle_{\mathcal{H}} = 0$$

$$\left\langle 0, \sum_{m=1}^{n} \frac{\alpha_m K(., y_m)}{n} - \sum_{m=1}^{n}\frac{\alpha_m}{n} \sum_{i=1}^{n} \frac{K(., x_i)}{n} \right\rangle_{\mathcal{H}} = 0$$

$$0 = 0$$

□

In homework 1, you have learned to sample functions from a RKHS with kernel K over the set $\mathcal{X} = \{-m, \ldots, m\}$ according to the probability

$$p(f) = Ce^{-\frac{||f||^2}{2}} \tag{3}$$

Here, you are asked to do the same thing but over the RKHS of a centered kernel K_c. Specifically, choose $m = 10$, $n = 11$, $x_1 = -10, x_2 = -9, \ldots, x_{11} = 0$.

$$\begin{aligned}
K_c(x,y) &= \left\langle K(., x) - \frac{1}{n}\sum_{i=1}^{n} K(., x_i), K(., y) - \frac{1}{n}\sum_{i=1}^{n} K(., x_i) \right\rangle \\
&= \left\langle K(., x), K(., y) \right\rangle + \left\langle K(., x), \frac{1}{n}\sum_{i=i}^{n} K(., x_i) \right\rangle + \\
&\quad \left\langle K(., y), \frac{1}{n}\sum_{i=1}^{n} K(., x_i) \right\rangle + \left\langle \frac{1}{n}\sum_{i=1}^{n} K(., x_i), \frac{1}{n}\sum_{j=1}^{n} K(., x_j) \right\rangle \\
&= K(x,y) + \frac{1}{n}\sum_{i=1}^{n} K(x, x_i) + \frac{1}{n}\sum_{i=1}^{n} K(y, x_i) + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n} K(x_i, x_j)
\end{aligned}$$

Ex. 1

$$K_c(x,y) = e^{\frac{-(x-y)^2}{2\tau}} - \frac{1}{11}\sum_{i=-10}^{0} e^{\frac{-(x-i)^2}{2\tau}} - \frac{1}{11}\sum_{i=-10}^{0} e^{\frac{-(y-i)^2}{2\tau}} + \frac{1}{121}\sum_{i=-10}^{0}\sum_{j=-10}^{0} e^{\frac{-(i-j)^2}{2\tau}}$$

Ex 2.

$$K_c(x,y) = (xy + 600)^2 - \frac{1}{11}\sum_{i=-10}^{0} (xi + 600)^2 - \frac{1}{11}\sum_{i=-10}^{0} (yi + 600)^2 + \frac{1}{121}\sum_{i=-10}^{0}\sum_{j=-10}^{0} (ij + 600)^2$$

Ex. 3

$$K_c(x,y) = tanh(xy+1) - \frac{1}{11}\sum_{i=-10}^{0} tanh(xi+1) - \frac{1}{11}\sum_{i=-10}^{0} tanh(yi+1) + \frac{1}{121}\sum_{i=-10}^{0}\sum_{j=-10}^{0} tanh(ij+1)$$

Ex. 4

$$K_c(x,y) = [xy+50] - \frac{1}{11}\sum_{i=-10}^{0}[xi+50] - \frac{1}{11}\sum_{i=-10}^{0}[yi+50] + \frac{1}{121}\sum_{i=-10}^{0}\sum_{j=-10}^{0}[ij+50]$$

3. (**) Write K_c in term of K using matrix operations.

$$\begin{aligned} K_c(x,y) &= K(x,y) + \frac{1}{n}\sum_{i=1}^{n}K(x,x_i) + \frac{1}{n}\sum_{i=1}^{n}K(y,x_i) + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}K(x_i,x_j) \\ &= K(x,y) - \frac{1}{n}\boldsymbol{K}_{\boldsymbol{y}}^T\mathbb{1} - \frac{1}{n}\boldsymbol{K}_{\boldsymbol{x}}^T\mathbb{1} + \frac{1}{n}\mathbb{1}^T\boldsymbol{K}\mathbb{1} \end{aligned}$$

where

$$\begin{aligned} \boldsymbol{K} &= [K(x_i,x_j)]_{ij\ (n,n)} \\ \boldsymbol{K}_{\boldsymbol{y}} &= [K(y,x_i)]_{i\ (1,n)} \\ \boldsymbol{K}_{\boldsymbol{y}} &= [K(x,x_i)]_{i\ (1,n)} \\ \mathbb{1} &= [1]_{i\ (1,n)} \end{aligned}$$

The centered kernel matrix can be written as:

$$\begin{aligned} \boldsymbol{K_c} &= \boldsymbol{K} - \boldsymbol{UK} - \boldsymbol{KU} - \boldsymbol{UKU} \\ &= (\boldsymbol{I} - \boldsymbol{U})\boldsymbol{K}(\boldsymbol{I} - \boldsymbol{U}) \end{aligned}$$

where

$$\boldsymbol{U} = \left[\frac{1}{n}\right]_{ij\ (n,n)}$$

$\boldsymbol{I}_{\ (n,n)}$ is the indentity matrix

4. Choose the Gaussian kernel with $\tau = 10$, and show 10 samples.

$$\frac{1}{11}\sum_{i=-10}^{0} f(i) = 0 \tag{4}$$

up to numerical errors.