Problem 1

Solution

Symmetric $K(x,y) = \langle K(x,\cdot), K(y,\cdot)\rangle = \overline{\langle K(y,\cdot), K(x,\cdot)\rangle} = \overline{K(y,x)} = K(y,x)$.

Positive semi-definite For any $n \in \mathbb{N}^+$ and $x_1,\ x_2,\ \cdots,\ x_n \in X$, construct a matrix M such that $M_{i,j} = K(x_i, x_j)$. For all $v \in \mathbb{R}^n$, we have $v^\intercal M v = \langle \sum_{i=1}^n v_i K(x_i,\cdot), \sum_{i=1}^n v_i K(x_i,\cdot)\rangle \geq 0$. Therefore, M is positive semi-definite.

Problem 2

Solution

(a) For all $x,\ y \in X$, construct a matrix $M = \begin{bmatrix} K(x,x) & K(x,y) \\ K(x,y) & K(y,y) \end{bmatrix}$. Then, M is positive semi-definite. Thus, $\det(M) = K(x,x)K(y,y) - K(x,y)^2 \geq 0$.

(b) (1) $\langle f,g\rangle = \sum_{i=1}^n \sum_{j=1}^m \alpha_i\beta_j K(x_i,x_j) = \sum_{i=1}^n \sum_{j=1}^m \alpha_i\beta_j K(x_j,x_i) = \langle g,f\rangle = \overline{\langle g,f\rangle}$; (2) For all $a,b \in \mathbb{R}$ and $f = \sum_{i=1}^n \alpha_i K(x_i,\cdot)$, $g = \sum_{i=1}^m \beta_i K(x_i,\cdot)$, $h = \sum_{i=1}^k \gamma_i K(x_i,\cdot)$, without loss of generality, suppose that $m \leq n$, and set $\beta_{m+1},\cdots,\beta_n = 0$. We have $\langle af+bg,h\rangle = \sum_{i=1}^n \sum_{j=1}^k (a\alpha_i + b\beta_i)\gamma_j K(x_i,x_j) = a\sum_{i=1}^n \sum_{j=1}^k \alpha_i\gamma_j K(x_i,x_j) + b\sum_{i=1}^m \sum_{j=1}^k \beta_i\gamma_j K(x_i,x_j) = a\langle f,h\rangle + b\langle g,h\rangle$; (3) For all $f = \sum_{i=1}^n \alpha_i K(x_i,\cdot)$, construct matrix M such that $M_{i,j} = K(x_i,x_j)$. Then, $\langle f,f\rangle = \alpha^\intercal M\alpha \geq 0$. Next, we prove $\langle f,f\rangle = 0 \iff f \equiv 0$. It is obvious that if $f \equiv 0$, $\langle f,f\rangle = 0$. If $\langle f,f\rangle = 0$, then for all $x \in X$, $|f(x)|^2 = \langle f, K(x,\cdot)\rangle^2 \leq \langle f,f\rangle\langle K(x,\cdot),K(x,\cdot)\rangle = 0$, and thus $f(x) = 0$. Note that here Cauchy-Schwarz inequality holds since it doesn't rely on the property we are currently proving. (I borrowed this result from page 37 in https://www.ism.ac.jp/~fukumizu/H20_kernel/Kernel_2_elements.pdf)

Problem 3

Solution

Let S be the subspace spanned by $K(x_1,\cdot),\cdots,K(x_n,\cdot)$. Then, for all f we can decomose it into $f = f_s + v$ such that $f_s \in S$ and $v \perp f_s$. We have $||f||^2 = ||f_s||^2 + ||v||^2 \geq ||f_s||^2$ and $g(||f_s||) \leq g(||f||)$. We also have, for all x_i, $f(x_i) = \langle f, K(x_i,\cdot)\rangle = \langle f_s, K(x_i,\cdot)\rangle + \langle v, K(x_i,\cdot)\rangle = \langle f_s, K(x_i,\cdot)\rangle = f_s(x_i)$. Thus, for a minimizer f^*, we must have $||v|| = 0$, $v = \vartheta$, and f^* can be linearly represented by $K(x_1,\cdot),\cdots,K(x_n,\cdot)$.