CS-E4830: Kernel methods of machine learning - Assignment 1

Christian Segercrantz 481056

March 16, 2022

Task 1

$$K_1(x, y) = (\langle x, y\rangle + c)^m \tag{1}$$

We know that the inner product $\langle x, y\rangle$ is a kernel. By the binomial theorem we get

$$\sum_{k=0}^{m} \binom{m}{k} \langle x, y\rangle^{m-k} c^k \tag{2}$$

Since we know that

1. A conic combination of kernels is a kernel
2. $c \geq 0$ and $\binom{m}{k} \geq 0$
3. The product of a kernel is a kernel

We can see that the polynomial kernel is a conic combination of the inner product to a power and thus a kernel:

$$\sum_{k=0}^{m} \binom{m}{k} \langle x, y\rangle^{m-k} c^k \tag{3}$$

$$= \sum_{k=0}^{m} C\langle x, y\rangle^{m-k}, \qquad C = \binom{m}{k} c^k \tag{4}$$

Additionally the last term, then the exponent is 0, is a constant. Hence we can re-write the sum as a sum of kernal and a constant:

$$K_{sum}(x, y) = \langle \phi(x), \phi(y)\rangle + C_m, \qquad C_m = \binom{m}{m} c^m = c^m \tag{5}$$

Hence K_1 is a kernel.

Task 2

$$h(x) = \text{sgn}\left(\sum_{i=1}^{n} \alpha_i k(x, x_i) + b\right), \quad k(x, x_i) = \langle \phi(x), \phi(x_i)\rangle \tag{6}$$

$$= sgn(||\phi(x) - c_-||^2 - ||\phi(x) - c_+||^2) \tag{7}$$

$$||\phi(x) - c_-||^2 = \langle \phi(x) - c_-, \phi(x) - c_-\rangle \tag{8}$$

$$= \langle \phi(x), \phi(x)\rangle - \langle \phi(x), c_-\rangle - \langle \phi(x), c_-\rangle + \langle c_-, c_-\rangle \tag{9}$$

$$||\phi(x) - c_+||^2 = \langle \phi(x) - c_+, \phi(x) - c_+\rangle \tag{10}$$

$$= \langle \phi(x), \phi(x)\rangle - \langle \phi(x), c_+\rangle - \langle \phi(x), c_+\rangle + \langle c_+, c_+\rangle \tag{11}$$

$$||\phi(x) - c_-||^2 - ||\phi(x) - c_+||^2 = \langle \phi(x), \phi(x)\rangle - \langle \phi(x), c_-\rangle - \langle \phi(x), c_-\rangle + \langle c_-, c_-\rangle \tag{12}$$

$$- \langle \phi(x), \phi(x)\rangle + \langle \phi(x), c_+\rangle + \langle \phi(x), c_+\rangle - \langle c_+, c_+\rangle \tag{13}$$

$$= -2\langle \phi(x), c_-\rangle + \langle c_-, c_-\rangle + 2\langle \phi(x), c_+\rangle - \langle c_+, c_+\rangle \tag{14}$$

$$= 2\langle \phi(x), c_+\rangle - 2\langle \phi(x), c_-\rangle + \langle c_-, c_-\rangle - \langle c_+, c_+\rangle \tag{15}$$

Since the sign-function is unaffected by a positive scalar multiplier we will divide the expression with 2.

The definition of c_- and c_+ are

$$c_- = \frac{1}{m_-} \sum_{i \in I^-} \phi(x_i) \tag{16}$$

$$c_+ = \frac{1}{m_+} \sum_{i \in I^+} \phi(x_i) \tag{17}$$

We can write the above inner products as

$$\langle \phi(x), c_- \rangle = \left\langle \phi(x), \frac{1}{m_-} \sum_{i \in I^-} \phi(x_i) \right\rangle \tag{18}$$

$$= \frac{1}{m_-} \sum_{i \in I^-} \langle \phi(x), \phi(x_i) \rangle \tag{19}$$

$$= \frac{1}{m_-} \sum_{i \in I^-} k(x, x_i) \tag{20}$$

And similarly we get for the + side

$$\langle \phi(x), c_+ \rangle = \frac{1}{m_+} \sum_{i \in I^+} k(x, x_i) \tag{21}$$

And for the two components containing only the centroids we get

$$\langle c_-, c_- \rangle = \left\langle \frac{1}{m_-} \sum_{i \in I^-} \phi(x_i), \frac{1}{m_-} \sum_{j \in I^-} \phi(x_j) \right\rangle \tag{22}$$

$$= \frac{1}{m_-} \sum_{i \in I^-} \frac{1}{m_-} \sum_{j \in I^-} \langle \phi(x_i), \phi(x_j) \rangle \tag{23}$$

$$= \frac{1}{m_-^2} \sum_{i \in I^-} \sum_{j \in I^-} k(x_i, x_j) \tag{24}$$

$$\langle c_+, c_+ \rangle = \frac{1}{m_+^2} \sum_{i \in I^+} \sum_{j \in I^+} k(x_i, x_j) \tag{25}$$

$$b = \frac{1}{2} \langle c_-, c_- \rangle - \frac{1}{2} \langle c_+, c_+ \rangle = \frac{1}{2m_-^2} \sum_{i \in I^-} \sum_{j \in I^-} k(x_i, x_j) - \frac{1}{2m_+^2} \sum_{i \in I^+} \sum_{j \in I^+} k(x_i, x_j) \tag{26}$$

The remaining terms then become

$$\langle \phi(x), c_+ \rangle - \langle \phi(x), c_- \rangle = \frac{1}{m_+} \sum_{i \in I^+} k(x, x_i) - \frac{1}{m_-} \sum_{i \in I^-} k(x, x_i) \tag{27}$$

From which we can see that if x is part of the positive cluster, α_i has to be $\frac{1}{m_+}$ and if part of the negative cluster $-\frac{1}{m_-}$. We can thus write Equation 15 as

$$\sum_{i=1}^{n} \alpha_i k(x, x_i) + b, \qquad \alpha_i = \begin{cases} \frac{1}{m_+}, & y = 1 \\ -\frac{1}{m_-}, & y = -1 \end{cases} \tag{28}$$

$$\implies h(x) = \operatorname{sgn}\left(\sum_{i=1}^{n} \alpha_i k(x, x_i) + b \right) \tag{29}$$

Task 3

$$K_2(x, y) = \cos(x + y) \tag{30}$$

We prove that K_2 is not a kernel by contradiction. Assume that K_2 is a kernel, i.e. it has to be positive definite. This means that any point should fulfill the criteria that

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j k(x_i, x_j) \geq 0 \tag{31}$$

for all $n \geq 1$, $\alpha \in \mathcal{R}^n$, $x \in \mathcal{X}^n$. Let's take the example $\alpha_i = \alpha_j = a \neq 0$ and $x_i = x_j = 2\pi$ This gives us $a^2 \cos(\frac{2\pi}{2}) = a^2 \cdot -1 = -a^2$. Since a is real and non-zero $-a^2 < 0$, which mean that the function is not positive definite. I.e. $\cos(x + y)$ is not a kernel!

Task 4

$$K_3(x, y) = \frac{1}{1 - xy}, \quad x, y \in (-1, 1) \tag{32}$$

Since $|xy| < 1, \frac{1}{1-xy}$ can be written as the series expansion

$$\frac{1}{1 - xy} = 1 + xy + x^2y^2 + x^3y^3 ... \tag{33}$$

Since products and conic sums of kernals are kernals and the sum of a kernel and a constant is a kernel as well can we conclude that $K_3(x, y)$ is a kernel.