XCS299i Problem Set #3

JUAN RICARDO PEDRAZA ESCOBAR

Machine Learning

Stanford Center for Professional Development

August,2021

Notation.

In the question, we use $K(*)$ to denote the kernel mapping, and use K to denote the corresponding kernel matrix $K_{j,k} = K(x^{(j)}, x^{(k)})$ generated by some collection $(x^{(i)})_{i=1}^{n} \subset R^d$.

1.a

$K(*)$ is a kernel. Let K_1 and K_2 denote the corresponding kernel metrics. By definition, the (j,k) term of the kernel matrix generated by K is

$$K_{j,k} = K(x^{(j)}, x^{(k)})$$
$$= K_1(x^{(j)}, x^{(k)}) + K_2(x^{(j)}, x^{(k)})$$
$$= K_{1,j,k} + K_{2,j,k}$$

Therefore,

$$K = K_1 + K_2.$$

Because K_1 and K_2 are symmetric and also $K_1(*)$ and $K_2(*)$ are valid kernels. Note that both K_1 and K_2 are positive-semi definite matrices, thus,

$$z^T K_Z = z^T K_{1Z} + z^T K_{2Z} \geq 0$$

So, $K(*)$ is a valid kernel. Because the corresponding kernel matrix generated by $K(*)$ is symmetric, for any $x^{(i)}$, and positive-semi definite.

1.b

$K(*)$ is not a kernel. A counter-example can be constructed as:

$$K_1(x, y) := 1\{x = y\}$$

$$K_2(x, y) := 2 \times 1\{x = y\}$$

Then for any $x^{(i)}$sequence, the corresponding kernel matrices are

$$K_1 = I_n$$

$$K_2 = 2I_n$$

We can say that $K_2 = 2K_1$ then,

$$z^T \ K_z = z^T \ (K_1 + 2\, K_1)z = -z^T K_1 z \ \leq 0$$

So, K is not positive-semi definite matrix. Hence, $K(*)$ is not a valid kernel.

1.c

$K(*)$ is a kernel.

For any $x^{(i)}$, let K_1 denote the kernel matrix generated by kernel $K_1(*)$, which is positive-semi definite matrix and symmetric. And the kernel matrix $K = aK_1$ is symmetric. And because for every $z \in R^n$,

$$z^T K_1 z \ \geq \ 0$$

Which implies

$$a\, z^T K_1 z \ \geq 0$$

So K is positive-semi definite matrix and $K(*)$ is a valid kernel

1.d

$K(*)$ is not a kernel. Let K and K_1 denote the corresponding kernel matrices. Then, K_1 is positive-semi definite matrix. Therefore, for every $z \in R^n$,

$$z^T K_1 z \geq 0$$

Then

$$(-a)\, z^T K_1 z \leq 0$$

$$z^T K\ z \leq 0$$

So, K is not a positive-semi definite matrix.

1.e

$K(*)$ is a Kernel. $K_1(*)$ and $K_2(*)$ are kernels, thus $\phi^{(1)}$ such that $K_1(\mathrm{x,z}) = \phi^{(1)}(x)^T\ \phi^{(1)}(z) = \sum_i \phi_i^{(1)}(x)\ \phi_i^{(1)}(z)$ and $\phi^{(2)}$ $K_2(\mathrm{x,z}) = \phi^{(2)}(x)^T\ \phi^{(2)}(z) = \sum_i \phi_j^{(2)}(x)\ \phi_j^{(2)}(z)$.

Therefore,

$$K(x,z) = K_1(x,z)\, K_2(x,z)$$

$$= \sum_i \phi_i^{(1)}(x)\phi_i^{(1)}(z) \quad \sum_j \phi_j^{(2)}(x)\ \phi_j^{(2)}(z)$$
$$= \sum_i \sum_j \left(\phi_i^{(1)}(x)\phi_j^{(2)}(x)\right)\left(\phi_i^{(1)}(z)\phi_j^{(2)}(z)\right)$$

Now we use $\psi_{(i,j)}(x) = \phi_i^{(1)}(x)\phi_j^{(2)}(x)$ and $\psi_{(i,j)}(z) = \phi_i^{(1)}(z)\phi_j^{(2)}(z)$ so,

$$= \sum_{(i,j)} \psi_{(i,j)}(x)\,\psi_{(i,j)}(z)$$

This shows us that we can express K as,

$$K(x,z) = \psi_{(i,j)}(x)^T\,\psi_{(i,j)}(z)$$

So, $K(*)$ is a Kernel.

1.f

$K(*)$ is a Kernel, and we have,

$$K_{i,j} = f\left(x^{(i)}\right)f\left(x^{(j)}\right) = f\left(x^{(j)}\right)f\left(x^{(i)}\right)$$

So, the generated kernel matrix is symmetric,

$$z^T K_1 z = \sum_{i=1}^{n}\sum_{j=1}^{n} z_i z_j K_{i,j}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} z_i f\left(x^{(i)}\right) z_j f\left(x^{(j)}\right)$$

$$\langle f\left(x^{(i)}\right)_{i=1}^{n}, z\rangle^2 \geq 0$$

Therefore, K is positive-semi definite matrix. So $K(*)$ is a valid kernel.

1.g

$K(*)$ is a valid kernel. Since K_3 is a kernel over $R^p \times R^p$, We are going to see that $K(*)$ is a valid kernel on $R^d \times R^d$ by showing it can be written as an inner product of feature mapping $\lambda \circ \phi : R^d \to R^l$ there exists a feature mapping $\lambda : R^p \to R^l$ for some l, such that $K_3\ (x, y) = \langle \lambda\ (x), \lambda(y)\rangle$ for every x, y $\in R^p$., then by definition

$$K(x, z) = K_3\left(\phi(x), \phi(z)\right)$$

$$= (\lambda(\phi(x)), \lambda\ (\phi(z)))$$

$$= (\lambda \circ \phi(x)), \lambda \circ \phi(z)))$$

Therefore $K(*)$ is a kernel.

1.h

- $K(*)$ is a valid kernel. Taking account, the result from (e) and induction, it can be shown that $K(x,z) = \prod Ki(x,z)$, where all $K_i(*)$ are kernels. This result permit said that $K'(x,z) = K(x,z)^n$ is also a kernel setting all $K_i(*)$.

- Let α_i denote the coefficients of p, and note also that $\alpha_i \geq 0$ for every i. Then by result from part (e) and part (c), $Ki(x,z) = \alpha_i K_1(x,z)^i$ is a valid kernel. And by result from part (a) induction, $K(x,z) = \sum_i K_i(x,z)$ is a valid kernel,