Exercise 4

1)

it is known that $e^{x} = \sum_{k=0}^{\infty} \frac{x^k}{k!}$, $-\infty < x < \infty$

$$K(x,y) = \exp(-\lambda (x-y)^2) = \exp(-\lambda x^2 - \lambda y^2 + 2\lambda xy)$$

$$= \exp(-\lambda x^2 - \lambda y^2) \cdot \exp(2\lambda xy)$$

$$= \exp(-\lambda x^2 - \lambda y^2) \cdot \left(\frac{2\lambda xy^0}{0!} + \frac{2\lambda xy}{1!} + \cdots + \frac{2\lambda xy^{\infty}}{\infty!} \right)$$

$$= \exp(-\lambda x^2 - \lambda y^2) \cdot \left(1 + \sqrt{\frac{2\lambda}{1!}} x \cdot \sqrt{\frac{2\lambda}{1!}} y + \cdots + \sqrt{\frac{(2\lambda)^{\infty}}{\infty!}} x^{\infty} \cdot \sqrt{\frac{(2\lambda)^{\infty}}{\infty!}} y^{\infty} \right)$$

let $\phi(x) = \exp(-\lambda x^2) \left[1, \sqrt{\frac{2\lambda}{1!}} x, \cdots, \sqrt{\frac{(2\lambda)^{\infty}}{\infty!}} x^{\infty} \right]^T$

then $\phi(y) = \exp(-\lambda y^2) \left[1, \sqrt{\frac{2\lambda}{1!}} y, \cdots, \sqrt{\frac{(2\lambda)^{\infty}}{\infty!}} y^{\infty} \right]^T$

Therefore, $K(x,y) = \phi(x)^T \phi(y)$.

I prefer to solve dual representation.

for primal question, we need solve $\min_{w,b} \max_{\alpha} L$

to find w that minize the $\max_{\alpha} L$, the solving complexity is relate to the dimesion of w. Applying kernel method, the dimesion of w is ∞. Thus, the solving complexity is really large.

for dual question, we need solve $\max_{\alpha} \min_{w,b} L$

to find α that maximize the $\min_{w,b} L$, the solving complexity is relate to the sample size of input X. Applying kernel method will not change the sample size. Thus, the solving complexity is small.

2) $\frac{1}{1-xy} = \sum_{n=0}^{\infty} (xy)^n = 1 + xy + (xy)^2 + \dots + (xy)^\infty$, $-1<xy<1$

let $\phi(x) = [1, x, x^2, \dots x^\infty]^T$, $\phi(y) = [1, y, y^2, \dots y^\infty]^T$

so $\phi(x)^T \phi(y) = \frac{1}{1-xy}$.

cause feature map exists and domain of xy is satisf

so it is PSD and log(1+xy) is a valid kernel.

3) $\log(1+xy) = xy - \frac{(xy)^2}{2} + \frac{(xy)^3}{3} - \dots + (-1)^{(n+1)} \frac{(xy)^n}{n}$, $-1<xy<1$

feature map does not exist cause the domain of xy conflict to domain of xy given by question.

so it is not PSD and log(1+xy) is not a valid kernel for this question.

4)

Assume a 2×1 vector $[x, y]^T = [1, \pi-1]$

then $M = \begin{bmatrix} \cos(2) & \cos(\pi) \\ \cos(\pi) & \cos(\pi-2) \end{bmatrix}$

for $X \in R^{2\times1}$, for example $[1, 1]^T$

$$[1, 1] \begin{bmatrix} \cos(2) & \cos(\pi) \\ \cos(\pi) & \cos(\pi-2) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= [-0.42-1, -1+0.42] \begin{bmatrix} 1 \\ 1 \end{bmatrix} = [-1.42, -0.58] \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -2 < 0$$

so it is not PSD, feature map do not exists