Introducción a ~~las~~ redes neuronales
CO-6612.

Eros Cedeño
16-10216.

Tarea 5

1. Para demostrar que $K(x,z) = e^{-\frac{\|x-z\|^2}{2\sigma^2}}$ es un Kernel se usaron los siguientes propiedades:

a) $K(x,z) = cK_1(x,z)$ con $c>0$ y K_1 un Kernel

b) $K(x,z) = f(x)K_1(x,z)f(z)$ con $f: \mathbb{R}^n \to \mathbb{R}$.

c) $K(x,z) = e^{K_1(x,z)}$ con K_1 un Kernel.

Se demostrará cada una de las propiedades. Sea S un conjunto finito de puntos $\{x_1, \dots, x_m\}$, y sea K_1 la matriz del kernel K_1 limitado a los puntos en S. Esto es:

$K_1 = [K_1(x_i, x_j)]$ con $x_i, x_j \in S$. Por definición, la matriz K_1 es semi-positiva definida, por lo tanto, $\forall \alpha \in \mathbb{R}^m$ se cumple que $\alpha' K \alpha \geq 0$. Ademas sea $f: \mathbb{R}^n \to \mathbb{R}$.

Demostración a): $\boxed{K(x,z) = cK_1(x,z)}$

Se tiene que K_1 es semipositiva definida, por lo tanto que $\alpha' K_1 \alpha \geq 0$, y como por hipótesis $c>0$ entonces $\alpha'(cK_1)\alpha = c(\alpha' K_1 \alpha) \geq 0$ lo que indica que la matriz cK_1 es semipositiva definida y en consecuencia $cK_1(x,z)$ es un Kernel.

Demostración b): $\boxed{K(x,z) = f(x)K_1(x,z)f(z)}$

Por definición de Kernel, existe una función característica tal que $K_1(x,z) = \langle \phi(x), \phi(z) \rangle = \sum_{i=1}^{t} \phi_i(x)\phi_i(z)$.

Por lo tanto

$$f(x)K_1(x,z)f(z) = f(x)\left(\sum_{i=1}^{t} \phi_i(x)\phi_i(z)\right)f(z)$$
$$= \sum_{i=1}^{t} f(x)\phi_i(x)\phi_i(z)f(z)$$

$$= \sum_{i=1}^{k} (f(x)\cdot\phi_i(x))(\phi_i(z)\cdot f(z))$$

$$= \sum_{i=1}^{k} (f(x)\phi(x))_i \cdot (f(z)\phi(z))_i$$

$$= \langle f(x)\phi(x), f(z)\phi(z) \rangle$$

. Sea $\varphi : x \longrightarrow f(x)\phi(x)$ luego

$K(x,z) = \langle \varphi(x), \varphi(z) \rangle$. que es la definición de Kernel. $\therefore$ K es un Kernel.

Proposición 1.1 $K(x,z) = K_1(x,z) + K_2(x,z)$

Sea $\mathbb{K}_1$ y $\mathbb{K}_2$ los matrices asociados a los Kernels K_1 y K_2 restringidos a S, luego por definición $\forall \alpha \in \mathbb{R}^m$, $\alpha' \mathbb{K}_1 \alpha \geq 0$ y $\alpha' \mathbb{K}_2 \alpha \geq 0$. Como los dimensiones son consistentes se pueden sumar ambos matrices y aplicando distributividad izquierda y derecha se tiene que $\alpha'(\mathbb{K}_1 + \mathbb{K}_2)\alpha \geq 0$. y por lo tanto. $\mathbb{K}_1 + \mathbb{K}_2$ es semipositiva definida y $K_1 + K_2$ es un Kernel.

Proposición 1.2. $K(x,z) = f(x)f(z)$

Considerese como función caracteristica $\phi : x \longrightarrow f(x) \in \mathbb{R}$. entonces $K(x,z)$ es su Kernel correspondiente.

Proposición 1.4. $K(x,z) = p(K_1(x,z))$ con $p(x)$ un polinomio con coeficientes positivos.

La forma general del polinomio será:

$$p(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_e x^e$$

eligiendo x como $K_1(x,z)$ se tiene:

$$P(K_1(x,z)) = \alpha_0 + \alpha_1 K_1(x,z) + \alpha_2 (K_1(x,z))^2 + \dots + \alpha_\ell (K_1(x,z))^\ell .$$

De la Proposición 1.2., si se toma $f(x)$ como constante. se tiene que α_0 es un Kernel. De la proposición 1.3 se tiene que $(K_1(x,z))^\ell$ es un Kernel $\forall \ell > 0$. De la demostración a) se demostró que $cK_1(x,z)$ es un Kernel, y finalmente. de la proposición 1.1 se tiene que la soma de Kernels es un Kernel. Por lo tanto $P(K_1(x,z))$ es un Kernel.

Proposición 1.3. $K(x,z) = K_1(x,z) K_2(x,z)$.

Consideremos de nuevo los matrices K_1 y K_2 como se definieron anteriormente. Sea

$K = K_1 \otimes K_2$ el producto tensorial de los dos matrices obtenido de reemplazar cada entrada de K_1 por K_2 multiplicada por esa entrada. El producto tensorial de dos matrices semipositivos definidos es semipositiva definida dado que los autovalores del producto son todos los pares de productos de los autovalores de los dos componentes. La matriz correspondiente a la función K_1K_2 también conocida como producto de Schur H de K_1 y K_2 es una submatriz principal de K definida por un conjunto de columnas y el mismo conjunto de filas.

Dado que para cualquier $\alpha \in \mathbb{R}^m$ existe un correspondiente $\alpha_1 \in \mathbb{R}^{\ell^2}$ tal que $\alpha' H \alpha = \alpha_1 K \alpha_1 \geq 0$. entonces H es semipositiva definida como se requería.

demostración c : $\boxed{K(x,z) = e^{K_1(x,z)}}$

Considerese el polinomio siguiente:

$$P_1(x) = 1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \cdots + \frac{1}{n!}x^n$$

Si se considera el limite con $n \to \infty$ se tendrá el siguiente polinomio:

$$P(x) = \lim_{n\to\infty} P_1(x) = \sum_{i=0}^{\infty} \frac{1}{i!}x^i$$

la serie obtenida es convergente a e^x. (Serie de taylor).

luego sustituyendo x por $K_1(x,z)$ y por la proposición 1.4

$$K(x,z) = P(K_1(x,z)) = \sum_{i=0}^{\infty} \frac{1}{i!}(K_1(x,z))^i = e^{K_1(x,z)}$$