Mercer's Theorem Contradiction.

Tuesday, August 16, 2022 10:25 AM

Define kernel as

$$K(x,t) = \langle \psi(x), \psi(t) \rangle \qquad (1)$$

Given ϕ_i as the i^{th} eigenfunction
Mercer's theorem states

$$K(x,t) = \sum_{i=1} \lambda_i \phi_i(x) \phi_i(t) \qquad (2)$$

Since ϕ is a function in RKHS $\rightarrow \phi(x) = \langle \phi, \psi(x) \rangle$
Therefore, Mercer's theorem implies

$$K(x,t) = \sum_{i=1} \lambda_i \; \psi(x)^T \phi_i \phi_i^T \psi(t)$$

$$= \psi(x)^T \left[\sum_{i=1} \lambda_i \; \phi_i \phi_i^T \right] \psi(t)$$

Since $K(x,t) = \psi(x)^T \psi(t)$, $K(x,t) \neq \psi(x)^T \left[\sum_{i=1} \lambda_i \phi \phi_i^T \right] \psi(t)$

This result suggests that mercer's theorem is wrong.

What happened?

Resolving the contradiction.

When Mercer's theorem states that $K(x,t) = \sum_{i=1} \lambda_i \phi_i(x) \phi_i(t)$
the eigenfunctions is defined as the orthogonal bases
from the integral operator $\mathcal{L}$ where

$$\mathcal{L}f(x) = \int K(x,t) f(t)\, d\mu(t)$$

The function space is defined on $L^2(X, \rho)$, that is, squared integrable functions and <u>NOT RKHS.</u>

Correspondingly, the eigenfunctions in mercer's theorem is in $L^2(X, \rho)$ NOT RKHS. Therefore, the reproducing property is not allowed, e.i.,

$$K(x,t) = \sum_{i=1} \lambda_i \phi_i(x) \phi_i(t) \neq \sum_{i=1} \lambda_i \langle \varphi(x), \phi_i \rangle \langle \varphi(t), \phi_i \rangle$$

The corresponding Operator that does reside in RKHS is defined as

$$T_n = \frac{1}{n} \sum_{j=1}^{N} \varphi(x_i) \varphi(x_i)^T$$

We denote its eigenfunctions as $\{v_1, v_2, v_3, \ldots\}$

According to Rosasco T_n has identical eigenvalues as $\mathcal{L}$

And their eigenvectors are related by

$$\phi_i = \frac{1}{\sqrt{\lambda_i}} v_i$$

Therefore, with mercer's theorem in RKHS, it should be

$$K(x,y) = \sum_{i=1} \lambda_i \left[\frac{1}{\sqrt{\lambda_i}} v_i(x)\right]\left[\frac{1}{\sqrt{\lambda_i}} v_i^T(y)\right]$$

We are now "allowed" to apply the reproducing property.

$$K(x,y) = \sum_{i=1} \frac{\lambda_i}{\lambda_i} v_i(x)\, v_i(y)^T$$

$$= \psi(x)^T \left[\sum_{i=1} v_i v_i^T\right] \psi(y)$$

← Since $\left[\sum_{i=1} v_i v_i^T\right]$ is a rotation matrix, the inner product is not impacted.

$$= \psi(x)^T \psi(y)$$

$$= \langle \psi(x), \psi(y) \rangle$$

$\square$