III. 3. Mercer's Theorem.

We begin with the current terminology belonging to this context. A kernel on a set X is a function $K : X \times X \to \mathbb{R}$; it is symmetric if $K(x,y) = K(y,x)$ $\forall x, y \in X$.

Def. III. 19 A symmetric kernel K on a set X is positive semidefinite if $\forall n \geq 1$ and $(x_1, .., x_n) \in X^n$, the symmetric matrix

$$\left(K(x_i, x_j) \right)_{i,j}$$

is positive semidefinite. That is $\forall (c_1, \ldots, c_n) \in \mathbb{R}^n$,

$$\sum_{i,j=1}^{n} c_i c_j K(x_i, x_j) \geqslant 0 .$$

The case $n=1$ implies

$$K(x,x) \geqslant 0 \quad \forall x \in X \qquad (\text{III}.20).$$

<u>Example III.21</u>: If $\mathcal{H}$ is a $\mathbb{R}$-Hilbert space and $\varphi : X \longrightarrow \mathcal{H}$ is any map, then

$$K(x,y) := \langle \varphi(x), \varphi(y) \rangle$$

is a positive semidefinite kernel on X.

In our context we will take (X,d) to be a compact metric space endowed with a regular Borel probability measure $\mu \in M^1(X)$. Given $K \in C(X \times X, \mathbb{R})$ continuous we know from Prop. 3.12 that the operator

$$T_K : L^2(X,\mu) \longrightarrow L^2(X,\mu),$$

$$T_K f(x) = \int_X K(x,y) f(y)\, d\mu(y) \quad \text{is}$$

Hilbert-Schmidt and hence compact.

If in addition K is a symmetric kernel, T_K is self-adjoint and the spectral theorem (Thm 3.14) applies. Observe that our hypothesis on X and μ garantees that $L^2(X,\mu)$ is separable.

Thm $\underbrace{\text{III}.22}$ (Mercer) Let (X,d) be a compact metric space, $\mu \in M^1(X)$ Borel regular such that $\forall U \subset X$ open non-empty $\mu(U) > 0$. Let $K \in C(X \times X)$ be a continuous positive semi-definite kernel on X.

Then there is an ONB $\{\varphi_1, \varphi_2, \dots\}$ of $(\operatorname{Ker} T_K)^\perp$ consisting of continuous eigenfunctions of T_K and if λ_i is the eigenvalue corresponding to φ_i,

$$\lambda_i > 0 \quad \forall i \geq 1.$$

In addition: $K(x,y) = \sum_{n=1}^{\infty} \lambda_n \varphi_n(x) \varphi_n(y)$

the sum being absolutely and uniformly convergent.

Observe that the fact that T_K is Hilbert–Schmidt gives us,

$$\sum_{n=1}^{\infty} \lambda_n^2 = \sum_{n,m=1}^{\infty} |\langle T_K \varphi_n, \varphi_m \rangle|^2 < +\infty .$$

While here:

<u>Cor. III.23</u> In the situation of Thm III.22 we have $K(x,x) = \sum_{n=1}^{\infty} \lambda_n \varphi_n(x)^2$,

with uniformly convergent right hand side.

In particular $\sum_{n=1}^{\infty} \lambda_n = \int_X K(x,x)\, d\mu(x) < +\infty$.

We preceed the Proof of Mercer's theorem with the following lemma whose proof is left as an exercise.

Lemma III. 24. Let (X, d) and $\mu \in M^1(X)$ be as in Mercer's theorem. In particular $\mu(U) > 0$ $\forall$ $U \subset X$ open, $\neq \phi$. Given $K \in C(X \times X)$ symmetric, the following are equivalent:

(1) K is positive semi-definite.

(2) $$\int_X \int_X f(x) f(y) K(x,y)\, d\mu(x)\, d\mu(y) \geq 0 \quad \forall f \in C(X)$$

(3) $\langle T_K f, f \rangle \geq 0$ $\forall f \in L^2(X, \mu)$.

Proof of Mercer's Thm.

(1) We start by observing that $\forall f \in L^2(X,\mu)$

$$T_K f(x) = \int_X K(x,y) f(y)\, d\mu(y) \text{ is well}$$

defined $\forall x \in X$ and continuous. Indeed

$$|T_K f(x_1) - T_K f(x_2)|$$

$$\leq \int_X |K(x_1,y) - K(x_2,y)|\, |f(y)|\, d\mu(y)$$

$$\leq \left[\int_X |K(x_1,y) - K(x_2,y)|^2 d\mu(y) \right]^{1/2} \|f\|_2$$

Now $K : X \times X \to \mathbb{R}$ being continuous on the compact metric space is uniformly continuous: in particular $\forall \varepsilon > 0 \;\exists \delta > 0$ such that $d(x_1,x_2) + d(y_1,y_2) < \delta$

$$|K(x_1,y_1) - K(x_2,y_2)| < \varepsilon.$$

In particular: $d(x_1, x_2) < \delta$

$\Rightarrow$ $|K(x_1,y) - K(x_2,y)| < \varepsilon \quad \forall y \in X$

which implies

$$|T_K f(x_1) - T_K f(x_2)| \leqslant \varepsilon \|f\|_2.$$

(2) By the spectral theorem let $f_1, f_2, \ldots, f_n, \ldots$ be an ONB of $(\operatorname{Ker} T_K)^\perp$ consisting of eigenvectors of T_K and λ_n the eigenvalue corr. to f_n. Then by lemma III.24:

$0 \leqslant \langle T_K f_n, f_n \rangle = \lambda_n$, and since $f_n \notin \operatorname{Ker} T_K$, we get $\lambda_n > 0$ $\forall n \geq 1$.

Thus we can write

$f_n = \dfrac{1}{\lambda_n} T_K f_n$, this being an equality in L^2! But by (1) we know that $T_K f_n$ is continuous; thus we define $\varphi_n(x) = \frac{1}{\lambda_n} T_K f_n(x)$ $\forall x \in X$; Then $\varphi_n \in C(X)$ and $\varphi_n = f_n$ almost everywhere. This proves the first part of the theorem.

(3) Now the fun begins!

Define $K_n(x,y) := K(x,y) - \sum_{k=1}^{n} \lambda_k \varphi_k(x) \varphi_k(y)$.

Then $K_n \in C(X \times X)$ and symmetric.

We claim that K_n is positive semi-definite.

Indeed let $f \in L^2(X,\mu)$, then

$$\langle T_{K_n} f, f \rangle = \langle T_K f, f \rangle - \sum_{k=1}^{n} \lambda_k \langle f, \varphi_k \rangle^2 .$$

Now expend f as follows:

$$f = \sum_{k=1}^{\infty} \langle f, \varphi_k \rangle \varphi_k + g$$

where g is the orthogonal projection of f on $\operatorname{Ker} T_K$. Then:

$$\langle T_K f, f \rangle = \sum_{k=1}^{\infty} \lambda_k \langle f, \varphi_k \rangle^2$$

and hence

$$\langle T_{K_n} f, f \rangle = \sum_{k=n+1}^{\infty} \lambda_k \langle f, \varphi_k \rangle^2 \geqslant 0.$$

The claim then follows from Lemma III.24.

In particular

$$K(x,x) - \sum_{k=1}^{n} \lambda_k \varphi_k(x)^2 = K_n(x,x) \geqslant 0$$

and hence $$\sum_{k=1}^{\infty} \lambda_k \varphi_k(x)^2 \leqslant K(x,x)$$

with left hand side (absolutely) convergent

convergent.

(4) Thus $\forall\ 1 \leq N \leq M$:

$$\sum_{k=N}^{M} \lambda_k |\varphi_k(x)| |\varphi_k(y)|$$

$$\leq \left(\sum_{k=N}^{M} \lambda_k \varphi_k(x)^2 \right)^{1/2} \underbrace{\left(\sum_{k=N}^{M} \lambda_k \varphi_k(y)^2 \right)}_{\leq K(y,y)}^{1/2}$$

$$\leq \left(\sum_{k=N}^{M} \lambda_k \varphi_k(x)^2 \right)^{1/2} \cdot \|K\|_b .$$

This implies that $\forall x \in X$

$$\sum_{k=1}^{\infty} \lambda_k \varphi_k(x) \varphi_k(y) \qquad (*)$$

converges absolutely and uniformly in y

and by symmetry $\forall y \in X$ $(*)$ converges absolutely and uniformly in x.

(8) Let now $K_x(y) = K(x,y)$, take any ONB $\psi_1, \dots, \psi_n, \dots$ of $\operatorname{Ker} T_K$ and expand $K_x \in L^2(X,\mu)$ in the ONB

$\varphi_1, \dots, \varphi_n, \dots, \psi_1, \dots, \psi_n$:

$$K_x = \sum_{k=1}^{\infty} \underbrace{(K_x, \varphi_k)}_{\lambda_n \varphi_n(x)} \varphi_k + \underbrace{\sum_{k=1}^{\infty} (K_x, \psi_k)\psi_k}_{\shortparallel \atop 0}$$

thus
$$K_x = \sum_{k=1}^{\infty} \lambda_k \varphi_k(x)\, \varphi_k .$$

This means $K_x = \lim_{n\to\infty} \sum_{k=1}^{n} \lambda_k \varphi_k(x)\varphi_k$

the convergence being in $L^2(X,\mu)$!

Thus there is a subsequence $(n_\ell)_{\ell \geq 1}$

such that $\displaystyle\sum_{k=1}^{n_\varepsilon} \lambda_k \varphi_k(x) \varphi_k(y)$ converges

almost everywhere in y to $K_x(y)$.

But this implies that $\forall x \in X$

the continuous functions

$$y \longmapsto \sum_{k=1}^{\infty} \lambda_k \varphi_k(x) \varphi_k(y)$$

and $y \longmapsto K(x,y)$ coincide almost

everywhere. Since $\mu(U) > 0$ $\forall U$ open, $\neq \emptyset$

this implies that they coincide everywhere.

Thus $$K(x,y) = \sum_{k=1}^{\infty} \lambda_k \varphi_k(x) \varphi_k(y) \quad \forall (x,y)$$

and in particular

$$K(x,x) = \sum_{k=1}^{\infty} \lambda_k \varphi_k(x)^2 \quad \forall x \in X.$$

(6) Now we show that the convergence of $\sum_{k=1}^{n} \lambda_k \varphi_k(x)^2$ to $K(x,x)$ is uniform.

Fix $\varepsilon > 0$ and let

$$V_n^{\varepsilon} = \left\{ x \in X : \sum_{k=1}^{n} \lambda_k \varphi_k(x)^2 > K(x,x) - \varepsilon \right\}.$$

Then V_n^{ε} is open $\forall n \geq 1$, ~~and~~ $V_n^{\varepsilon} \subset V_{n+1}^{\varepsilon}$ $\forall n \geq 1$ and by pointwise convergence,

$$\bigcup_{n \geq 1} V_n^{\varepsilon} = X.$$

Thus since X is compact we deduce $\forall \varepsilon > 0$ $\exists\, n(\varepsilon)$ with $V_{n(\varepsilon)}^{\varepsilon} = X$

which shows that the convergence is uniform.

(7) Going back to the inequality in (4):

$$\sum_{k=N}^{M} \lambda_k |\varphi_k(x)||\varphi_k(y)|$$

$$\leq \left(\sum_{k=N}^{M} \lambda_k \varphi_k(x)^2\right)^{1/2} \left(\sum_{k=N}^{M} \lambda_k (|\varphi_k(y)|)^2\right)^{1/2}$$

We deduce that $\displaystyle\sum_{k=1}^{\infty} \lambda_k \varphi_k(x) \varphi_k(y)$

converges absolutely and uniformly in $X \times X$

with sum $= K(x, y)$. $\boxed{\xi}$