Prove that K satisfies condition of Mercer's theorem

http://en.wikipedia.org/wiki/Mercer%27s_theorem

How to show that given $K(x_1, x_2)$ $\exists \phi \in \mathcal{H}$ s.t. $K(x_1, x_2) = \phi^T(x_1)\phi(x_2)$

Route 2

Route 1

If K is a p.s.d kernel then $\exists \phi$

$\forall n$ & $\forall x_1 \ldots x_n$, Gram matrix

$$\begin{bmatrix} K(x_1, x_1) & \cdots & K(x_1, x_n) \\ & & \vdots \\ \cdots & \cdots & K(x_n, x_n) \end{bmatrix}$$

is p.s.d

Properties of p.s.d kernels:

① $n=1$, $K(x_1, x_1)$ to be p.s.d matrix

i.e $K(x_1, x_1) \geq 0$

② $n=2$, $\begin{bmatrix} K(x_1, x_1) & K(x_1, x_2) \\ K(x_2, x_1) & K(x_2, x_2) \end{bmatrix}$

should be p.s.d

$\Rightarrow$ all eigenvalues are ≥ 0

$\Rightarrow$ sum & prod of eigenvalues ≥ 0

$\Rightarrow$ trace & determinant ≥ 0

$\Rightarrow$ $K(x_1, x_1) + K(x_2, x_2) \geq 0$

$K(x_1, x_1) K(x_2, x_2) - K(x_1, x_2)^2 \geq 0$

$\Downarrow$

$K(x_1, x_2)^2 \leq K(x_1, x_1) K(x_2, x_2)$

RKHS: Reproducing Kernel Hilbert space

<u>Claim</u>: The mapping:

$f(\cdot) : \mathcal{X} \to \mathcal{H}$

Like my $\phi(x)$

& corresponding $\langle \, , \, \rangle$ will give us back the kernel

$$\mathcal{H} = \left\{ f(\cdot) = \sum_{i=1}^{m} \alpha_i K(\cdot, x_i) \right\}$$

Assume α_i & K to be real valued for time being

Where $x_1, x_2 \ldots x_m \in \mathcal{X}$

$$\langle f, g \rangle = \sum_i \sum_j \alpha_i \beta_j K(x_i', x_j)$$

where $f = \sum_{i=1}^{m} \alpha_i K(\,, x_i')$

$g = \sum_{j=1}^{n} \beta_j K(\,, x_j)$

Claims: A $\langle f, g \rangle$ is well defined, i.e independent of choice of α_i, β_j & $i = 1..n$ & $j = 1..m$

(Proof)

$$\langle f, g \rangle = \sum_i \alpha_i \left(\sum_j \beta_j K(x_i', x_j) \right)$$

$$= \sum_i \alpha_i g(x_i') = \sum_j \beta_j f(x_j)$$

of independent β_j & x_j is

of independent α_i & x_i

Next prove: (Home work)

(B) $\langle f, g\rangle = \langle g, f\rangle$

(C) $\langle f_1 + f_2, g_1 + g_2\rangle = \langle f_1, g_1\rangle + \langle f_1, g_2\rangle + \langle f_2, g_1\rangle + \langle f_2, g_2\rangle$

Tricky:

(D) If $\langle f, f\rangle = 0$ then $f = 0$

(will require Cauchy Schwarz inequality we proved for K psd kernel)

(Assume we can next "complete" this inner product space to a Hilbert space)

Q: What is $\phi(x)$?

H/w: Prove

Ans: $\phi(x) = K(\cdot, x)$

$K(x_1, x_2) = \langle K(\cdot, x_1), K(\cdot, x_2) \rangle$