Pattern Recognition Machine Learning

vamshi, SC15M045

February 17, 2016

1

(a)

An n dimensional Euclidean space with norm defined as $\|x\|_p = (\sum |x_i|^p)^{\frac{1}{p}}$ is a normed space but not inner product space except for p=2, because this norm doesn't satisfy the **Parallelogram Equality** required of a norm to have an inner producy associated with it.
Parallelogram Equality:
In normed space, the statement of parallelogram law is an equation relating norms

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

For any norm satisfying the parallelogram law, the inner product generating the norm is unique and is given by

$$< x, y >= \frac{\|x + y\|^2 - \|x - y\|^2}{4}$$

(b)

$f, g \in C[0, 1], f(x) = x + 1, g(x) = sin(x)$
$< f, g >= \int_a^b f(x).g(x)dx$
$= \int_0^1 (x + 1).sin(x)dx$
$= [-(x + 1)cos(x) + sin(x)]_0^1$

2

M be a closed subspace of Hilbert space H(complete inner product space). $M \subset H$
let $\{X_n\}$ be the cauchy sequence in M $\implies$ $\{X_n\}$ be the cauchy sequence in H.
Since H is complete $X_n \longrightarrow x \in H$. x is the limit point of M.
since M is closed, $x \in M$. i.e, Every cauchy sequence in M, converges in M.
$\therefore$ M is complete $\implies$ M is Hilbert space.

3

Training points $(x_i, y_i), i = 1, 2, ...N$. Let $f \in F(RKHS)$ generates the data.f is entirely contained in $span\{k_{x_i}\}$. let Y = $span\{k_{x_i}\}$. As every finite dimensional subspace of a normed space X is closed in X, Y is closed.Therefore by projection theorem,

$$F = Y \oplus Y^{\perp}$$

Hence $f = f_y + f_{y^\perp}, f_y \in Y, f_{y^\perp} \in Y^\perp$. Now $f(x_i) = (< f, K_{x_i} >) = (< f_y, k_{x_i}) >$. As $f_y \in Y, f_y = \sum_{i=1}^{i=n} \alpha_i k_{x_i}$. Thus any function of the form $f_y = \sum_{i=1}^{i=n} \alpha_i k_{x_i} + f', f' \in Y^\perp$ satisfies the given points.

4

(a)

$f(x) = 3x_1 + 4x_2 + 5x_3, x = (x_1, x_2, x_3)$ $f(x) = w^T x, w = w = (3, 4, 5)^T$;.

$$\|f\| = \|w\| = 5\sqrt{2}$$

(b)

Let $M = \{k_{x_i}\}$ and let $f \in M^\perp$. Therefore $< f, k_x >= 0, \forall x \in X$. Therefore f(x) $= 0$ $\forall x \in X$. Hence f $\equiv$ 0.Hence $M^\perp = \{0\}$. Hence $\overline{span(M)} = F$. Therefore every $f \in F$ can be represented as

$$f = \sum \alpha_i k_{x_i}, \alpha_i \in \mathbb{R}$$

Hence $f(x) = \sum \alpha_i k_{x_i}(x) = \sum \alpha_i k(x_i, x)$. Therefore every element in RKHS can be represented as linear combination of k_{x_i}

(c)

(i)

$\tilde{f}(x) = f(x) + b, f \in \mathbb{F}$, Reproducing kernel $k(x, y) = < x, y >^2$

$$\begin{aligned} k(x,y) &= (x_1 y_1 + x_2 y_2)^2 \\ &= \begin{bmatrix} x_1^2 \\ x_2^2 \\ \sqrt{2} x_1 x_2 \end{bmatrix} \cdot \begin{bmatrix} y_1^2 \\ y_2^2 \\ \sqrt{2} y_1 y_2 \end{bmatrix} = \; < \phi(x), \phi(y) > \end{aligned}$$

$k(x, x) = 100, k(x, y) = k(y, x) = 9, k(y, y) = 1$. any $f \in \mathbb{F}$ can be represented as $f = \sum_{i=1}^{i=1} k_{x_i}$. Therefore

$$\begin{aligned} \|f\|^2 &= (< f, f >) = (< \alpha_1 k_{x_1} + \alpha_2 k_{x_2}, \alpha_1 k_{x_1} + \alpha_2 k_{x_2} >) \\ &= \alpha_1^2 k(x, x) + 2\alpha_1 \alpha_2 k(x, y) + \alpha_2^2 k(y, y) \\ &= 100\alpha_1^2 + 18\alpha_1 \alpha_2 + \alpha_2^2 \end{aligned}$$

unknown parameters are α_1, α_2, b

(ii)

$\tilde{f}$ is in the linear span of k_{x_i} only if b=0, but f is in the linear span of k_{x_i}. Kernel methods finds a unique f and unique b such that,it minimizes the regularized risk function and f has the minimum norm of all functions that can be solutions.