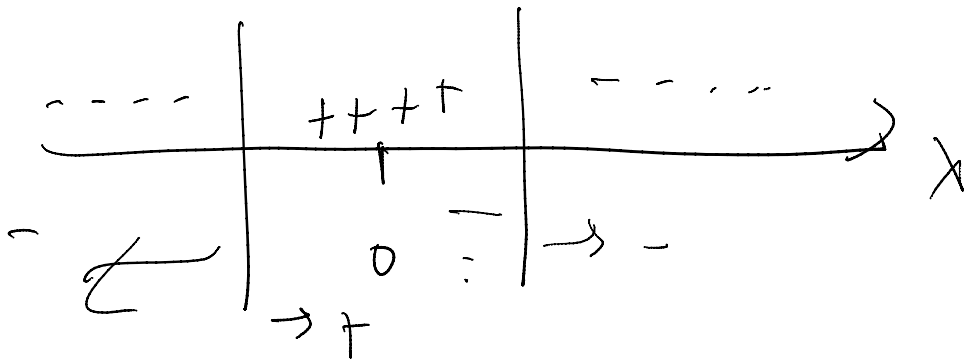
Lec 9: Tuesday, 29th October - Kernels (on features & labels)

29 October 2013
12:10



$$f(x) = \frac{1}{2} x^t \mu x + \Delta^t x + b$$

$x \to z$ $\quad z = ?\, x x^t$

① $f - D^2$, $\bar{D}$

② $x \to z$, D^2, nD^2, $O(n\bar{D})$, $\hat{D}$

D

$f(x) = a x^{20}$ ... $= D^{20}$

$n^2 D$ $\quad D^{20}$

MAX $1^t \alpha - \frac{1}{2} \alpha^t Y K Y \alpha$

α $\quad 0 < \alpha \leq C$, $1^t y \alpha = 0$

$K_{n \times n}$, $K_{ij} = \phi^t(x_i)\phi(x_j)$ 1

$O(n^3)$ – SMO

$K_{n \times n}$ – n^2 D

$\phi(x) =$

$\phi^t(x_i)\phi(x_j) = (x_i^t x_j + 1)^{20}$

n^2 1D

$K_{ij} = \phi^t(x_i)\phi(x_j)$

$w^t \phi(x)$, $w = X^t \alpha$ (D×N)

$\alpha^t y \phi^t(x_{i-n}) \phi(x) = \phi(x)^t y \alpha$ (D×N)

$$= \sum_{i=1}^{N} \alpha_i y_i K(x_i, x)$$

D

KERNELS

$K(x_i, x_j) =$ SIMILARITY

(1) LIN $K = x_i^t x_j = x_i^t \Sigma^{-1} x_j$

(2) POLY $= (x_i^t x_j + C)^p$

$-\gamma \|x_i - x_j\|_2^2$

③ RBF $k = e^{-\chi^2(x_i, x_j)}$

④ $\chi^2 = e$

⑤ ~~TANH~~, $c = \tanh(x_i^t x_j - c)$

MERCER

$x \in \mathbb{Z}$, y

$$\int\int_{x \in \mathbb{Z},\ y \in \mathbb{Z}} f(x)\, k(x,y)\, f(y)\, dx\, dy \geq 0$$

$\equiv$

$$X = K \geq 0$$

$$\left(K + \lambda_{\min} I \right)$$

($K + \lambda_{min}$)

RBF $k(x_i, x_0) = e^{-\gamma \|x_i - x_0\|_2^2}$

$\phi \to \infty$

$k(x_i, x_i) = 1 \geq 0$

$\gamma = 0 \Rightarrow k = 1$

$1 + \gamma - \frac{1}{2}\gamma^2 + \gamma^3$

$K_{ij} = 0 \quad i \neq j = 1$

KERNELS on LABELS

& STRUCTURED OUTPUT

$$f_w(x, y)$$

$$\min_{w,\xi} \tfrac{1}{2}\|f\|_2^2 + \sum_{i=1}^{n} \xi_i \qquad \forall i, y$$

$$f(x_i, y_i) \geq -(x_i, y) + \underbrace{\Delta(y_i, y)} - \xi_i$$

$$\xi \geq 0$$

$$f(x, y) = w^T \phi(x, y)$$

$$K(x_i, x_j, y_i, y_j) = \phi^T(x_i, y_i)\phi(x_j, y_j)$$

$$= k(x_i, x_j)\, k(y_i, y_j)$$

$$y^* = \arg\max_{y} f(x, y)$$