Date : 29-08-2018

SVM

(12)

Check if kernels are valid ?

Let $K(X, Z)$ be a kernel where $X, Z \in \mathbb{R}^n$ and are feature vectors. i.e. $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n \times 1}$ $Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}_{n \times 1}$

$K(X, Z)$ must be expressed in

$$\phi^T(x)\,\phi(z) \quad \text{or} \quad \langle \phi(x), \phi(z) \rangle$$

i.e. $K(X, Z) = \phi^T(x) \cdot \phi(z) = \langle \phi(x), \phi(z) \rangle$

$K(X, Z) \in \mathbb{R}$

Tutorial Problems.

Polynomial kernel

1) Check if $K(X, Z) = (X^T Z)^2$ is a valid kernel ?

Soln. Let $X, Z \in \mathbb{R}^n$

$$K(X, Z) = \left([x_1, x_2\ x_3 \cdots x_n] \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \right)^2$$

$$= \left(\sum_{i=1}^{n} x_i z_i \right)^2 = \left(\sum_{i=1}^{n} x_i z_i \right) \left(\sum_{j=1}^{n} x_i z_i \right)$$

$$= (x_1 z_1 + x_2 z_2 + \cdots + x_n z_n)(x_1 z_1 + \cdots + x_n z_n)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i z_i x_j z_j = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j z_i z_j$$

$$\phi(x) = \begin{bmatrix} x_1 x_1 \\ x_1 x_2 \\ x_1 x_3 \\ x_2 x_1 \\ \vdots \\ x_3 x_3 \end{bmatrix} \quad \phi(z) = \begin{bmatrix} z_1 z_1 \\ z_1 z_2 \\ z_1 z_3 \\ z_2 z_1 \\ \vdots \\ z_3 z_3 \end{bmatrix} \quad K(x, z) = \phi^T(x)\,\phi(z)$$

2. check if $K(x,z) = (x^T z + c)^2$ is valid kernel ?
$x, z \in \mathbb{R}^n$

$$K(x,z) = (x^T z + c)^2$$

$$= (x^T z)^2 + c^2 + 2c\, x^T z$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j z_i z_j + c^2 + 2c \sum_{i=1}^{n} x_i z_i$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j z_i z_j + c \cdot c + \cancel{\sqrt{2c}\,x\,\sqrt{2c}\,z}$$

$$\sum_{i=1}^{m} \sqrt{2c}\,x_i \; \sqrt{2c}\,z_i$$

$$\phi(x) = \begin{bmatrix} x_1 x_1 \\ x_1 x_2 \\ x_1 x_3 \\ \vdots \\ x_3 x_3 \\ \sqrt{2c}\,x_1 \\ \sqrt{2c}\,x_2 \\ \sqrt{2c}\,x_3 \\ c \end{bmatrix}$$

$$= \phi^T(x)\,\phi(z) + c \cdot c + \sum_{i=1}^{m} \sqrt{2c}\,x_i \; \sqrt{2c}\,z_i$$

Prove Gaussian kernel $K(x,z) = \exp\left(-\frac{\|x-z\|^2}{2\sigma^2}\right)$ is a valid kernel ?

Let $x, z \in \mathbb{R}^n$

$$K(x,z) = \exp\left(-\frac{\|x-z\|^2}{2\sigma^2}\right)$$

using $e^{a+b} = e^a \cdot e^b$

$$K(x,z) = \exp\left(-\frac{(\|x\|^2 + \|z\|^2 - 2\|x\|\|z\|)}{2\sigma^2}\right)$$

$$= \exp\left(-\frac{(x^T x + z^T z - 2x^T z)}{2\sigma^2}\right)$$

$$= e^{-\frac{x^T x}{2\sigma^2}} \cdot e^{\frac{-z^T z}{2\sigma^2}} \cdot e^{\frac{x^T z}{\sigma^2}}$$

it can be assamed as linear kernels, scaled by $\frac{1}{2\sigma^2}$ and exporent applied.

Date 28.5.2018

"Mercer theorem" for Validating kernels ?

Statement : Let $K(x,z)$ be given. Then K is a valid (mercer) kernel (i.e $\exists \phi$ s.t $K(x,z) = \phi^T(x)\,\phi(z)$. if and only if $\forall$ for all $\{x^{(1)}, x^{(2)}, \ldots x^{(m)}\}$ $m < \infty$ the kernel matrix $K \in \mathbb{R}^{m \times m}$ is symmetric positive semi definite.

To Prove ~~that~~ valid kernel matrix is positive-semi definite.

Proof : Let $K(x,z)$ be a valid kernel.

s.t $\exists \phi$ s.t $K(x,z) = \phi^T(x)\,\phi(z) = \langle \phi(x), \phi(z) \rangle$

Let set of points i.e training examples be $\{x^{(1)}, x^{(2)}, \ldots, x^{(m)}\}$

for any pair of points/ training examples $x^{(i)}, x^{(j)}$ we can define $K(x^{(i)}, x^{(j)})$. so, we can have such $m \times m$ kernels

$$K = \begin{bmatrix} K(x^{(1)}, x^{(1)}) & K_{12} & \cdots & K_{1m} \\ & & K_{ij} & \\ K_{m1} & K_{m2} & \cdots & K_{mm} \end{bmatrix}_{m \times m}$$

$$\underset{\substack{\uparrow \\ \text{Kernel matrix}}}{K} = \begin{bmatrix} K(x^{(1)}, x^{(1)}), & K(x^{(1)}, x^{(2)}) & \cdots & K(x^{(1)}, x^{(m)}) \\ \vdots & & & \vdots \\ K(x^{(m)}, x^{(1)}), & K(x^{(m)}, x^{(2)}) & \cdots & K(x^{(m)}, x^{(m)}) \end{bmatrix}_{m \times m}$$

$$\underset{\substack{\uparrow \\ ij^{th}\text{ element}}}{K_{ij}} = K(x^{(i)}, x^{(j)}) = \langle \phi(x^{(i)}), \phi(x^{(j)}) \rangle = \phi^T(x^{(i)})\,\phi(x^{(j)})$$

define some $z \in \mathbb{R}^m$.

write: $z^T K z$

$$z^T K z = \sum_{i=1}^{m}\sum_{j=1}^{m} z_i K_{ij} z_j$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{m} z_i \, \phi(x^{(i)})^T \phi(x^{(j)}) z_j$$

$$= \sum_{i=1}^{m}\sum_{j=1}^{m} z_i \sum_{k=1}^{m} \phi(x^{(i)})_k \, \phi(x^{(j)})_k z_j$$

$$= \sum_{k=1}^{m}\sum_{i=1}^{m}\sum_{j=1}^{m} z_i (\phi(x^{(i)}))_k \, (\phi(x^{(j)}))_k z_j$$

$$= \sum_{k=1}^{m}\sum_{i=1}^{m} \left(z_i (\phi(x^{(i)}))_k \right)^2 \geq 0$$

Sum of squares $\therefore$ $K \geq 0$

Positive semi definite.