Kernel Trick

Replace all inner products of x and x' (i.e., $\langle x, x' \rangle$) with Kernel function $K(x, x')$.

Kernelized Nearest Neighbor Classification

Two vectors x_1 and x_2

The Euclidian distance between x_1 & x_2:

$$\text{dist}(x_1, x_2) = \sqrt{(x_{11} - x_{21})^2 + (x_{12} - x_{22})^2 + \cdots + (x_{1D} - x_{2D})^2}$$

$$\text{dist}(x_1, x_2) = \text{dist}(x_2, x_1) = \underbrace{\|x_1 - x_2\|_2}_{l_2 \text{ norm of } x_1 - x_2}$$

$$\|x_1 - x_2\|_2^2 = (x_{11} - x_{21})^2 + (x_{12} - x_{22})^2 + \cdots + (x_{1D} - x_{2D})^2$$

For simplicity, let's assume x_1 & x_2 are 2-dimensional

$$\|x_1 - x_2\|_2^2 = (x_{11} - x_{21})^2 + (x_{12} - x_{22})^2$$

$$= \underline{(x_{11})^2} + (x_{21})^2 - 2x_{11}x_{21} + \underline{(x_{12})^2} + (x_{22})^2 - 2x_{12}x_{22}$$

$$x_1 = \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix}$$

$$x_2 = \begin{bmatrix} x_{21} \\ x_{22} \end{bmatrix}$$

$$x_1^T \cdot x_1 = \begin{bmatrix} x_{11} & x_{12} \end{bmatrix} \cdot \begin{bmatrix} x_{11} \\ x_{12} \end{bmatrix} = (x_{11})^2 + (x_{12})^2$$

$$= \langle x_{11}, x_{7} \rangle$$

$$\|x_1 - x_2\|_2^2 = \langle x_1, x_1\rangle + \langle x_2, x_2\rangle - 2\langle x_1, x_2\rangle$$
$$= K(x_1, x_1) + K(x_2, x_2) - 2K(x_1, x_2)$$

If we replace Euclidian distance with all Kernel functions such that

$$dist(x_1, x_2) = \sqrt{K(x_1, x_1) + K(x_2, x_2) - 2K(x_1, x_2)}$$

we end up with the <u>Kernelized Nearest Neighbor Classification.</u>

Kernelijed Lidge Regression.

$$w = X^T (XX^T + \lambda I_N)^{-1} \cdot Y$$

Assume for simplicity $N = 3$ $D = 2$.

Then X is 3×2

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix} \qquad X^T = \begin{bmatrix} x_{11} & x_{21} & x_{31} \\ x_{12} & x_{22} & x_{32} \end{bmatrix}$$

3×2 $\qquad$ 2×3

$$
XY^T = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix} \cdot \begin{bmatrix} x_{11} & x_{21} & x_{31} \\ x_{12} & x_{22} & x_{32} \end{bmatrix}
$$

$$
= \underbrace{\begin{bmatrix} (x_{11})^2 + (x_{12})^2 & x_{11}x_{21} + x_{12}x_{22} & x_{11}x_{31} + x_{12}x_{32} \\ x_{21}x_{11} + x_{22}x_{12} & \vdots & \\ x_{31}x_{11} + x_{32}x_{12} & & (x_{31})^2 + (x_{32})^2 \end{bmatrix}}_{3\times 3 \text{ gram matrix } K}
$$

$$\underbrace{\alpha}_{N\times 1} \triangleq \underbrace{(K + \lambda I_N)^{-1}}_{N\times N} \cdot \underbrace{y}_{N\times 1}$$

$$\underbrace{w}_{D\times 1} = \underbrace{X^T}_{D\times N} \cdot \underbrace{\alpha}_{N\times 1}$$