

$$L_D(w, b, \alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{m=1}^N \sum_{n=1}^M \alpha_m \alpha_n (X_m^T X_n) y_m y_n$$

$$\sum_{i=1}^N \alpha_i y_i = 0$$

def trainSVM (  $y_1 \dots y_N, x_1 \dots x_N$  )

figure out the best  $\alpha_1, \dots, \alpha_N$

and  $b$

and  $w = \sum_{i=1}^N \alpha_i y_i x_i$

return  ~~$\alpha_1, \dots, \alpha_N$~~ ,  $b, w$

Since every  $x_m$  occurs as a dot product with another  $x_n$ .

def trainSVM (  $y_1 \dots y_N, K$  )

where  $K$  is an  $N \times N$  matrix

and  $K_{mn} = K[m][n] = x_m^T x_n$

def testSVM (  $w, b, x^*$  )

if  $w^T x^* + b \geq 0$  return +1

$< 0$  return -1

$$w^T x^* = \left( \sum_{i=1}^N \alpha_i y_i x_i \right)^T x^*$$

$$= \left[ \sum_{i=1}^N \alpha_i y_i (x_i^T x^*) \right]$$

Let  $k^*$  be a  $N \times 1$  vector:

$$k^*[i] = x_i^T x^*$$

def testSVM ( $\alpha_1, \dots, \alpha_N, b, k^*$ )

$$if \left( \sum_{i=1}^N \alpha_i y_i k^*[i] \right) + b \geq 0 \text{ return } +1$$

$$< 0 \text{ return } -1$$

$$K[m][n] = x_m^T x_n$$

→ Some similarity between  $x_m$  &  $x_n$

I can replace  $(x_m^T x_n)$  with some other function,

$$k(x_m, x_n)$$

→ Kernel SVM

Option 1 to produce a non-linear SVM

Choose some basis fn.

Replace  $x_m \rightarrow \phi_1(x_m), \phi_2(x_m), \dots$

Then feed this to the standard SVM.

Option 2

To use a different  $k(x_m, x_n)$  instead

$$of \underline{x_m^T x_n}$$

$$\begin{matrix} x_m & x_n \\ \uparrow & \nearrow \end{matrix}$$

Two data instances

$$k(x_m, x_n) = \mathbb{R}$$

A valid kernel should be such that:  
There should be a basis function expansion, such that

$$\begin{matrix} \Phi_m & \begin{bmatrix} \phi_1(x_m) \\ \phi_2(x_m) \\ \phi_3(x_m) \\ \vdots \\ \phi_p(x_m) \end{bmatrix} \end{matrix} \quad \begin{matrix} \phi_1 \phi_2 \phi_3 \dots \phi_p \\ \Phi_n \begin{bmatrix} \phi_1(x_n) \\ \phi_2(x_n) \\ \vdots \\ \phi_p(x_n) \end{bmatrix} \end{matrix}$$

$$\Phi_m^T \Phi_n = k(x_m, x_n)$$

If  $\Phi$  is identity  $\Rightarrow \Phi(x_m) = x_m$

$$k(x_m, x_n) = (1 + x_m^T x_n)^2$$

What will be  $\Phi$ ?

Assume  $x_m \in \mathbb{R}^2$   
 $x_m = [x_{m1}, x_{m2}]$

$$\begin{bmatrix} 1 \\ x_{m1} \\ x_{m2} \\ \sqrt{2}x_{m1} \\ \sqrt{2}x_{m2} \\ \sqrt{2}x_{m1}x_{m2} \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ x_{m1}^2 \\ x_{m2}^2 \\ \sqrt{2}x_{m1} \\ \sqrt{2}x_{m2} \\ \sqrt{2}x_{m1}x_{m2} \end{bmatrix}$$