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

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

def trainSVM($y_{1...N}$, $x_{1...N}$)

figure out the best $\underline{\alpha_1, \ldots \alpha_N}$

and b

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

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

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

def trainSVM($y_{1...N}$, 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

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

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

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

$$\boxed{k^*[i] = x_i^T x^*}$$

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

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

< 0 return −1

$$\boxed{K[m][n] = x_m^T x_n} \rightarrow$$ Some similarity between x_m & x_n

I can replace $(x_m^T x_n)$ with some other function, $\underline{k(x_m, x_n)}$.

→ Kernel SVM

Option 1 to produce a non-linear SVM

Choose some basis fn.

Replace $[x_m \longrightarrow \phi_1(x_m), \phi_2(x_m) \ldots]$

Then feed this to the standard SVM.

Option 2

To use a different $k(x_m, x_n)$ instead of $\underline{x_m^T x_n}$

$x_m \quad x_n$

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

$$\phi_1 \; \phi_2 \; \phi_3 \cdots \phi_p$$

$$\phi_m \begin{bmatrix} \phi_1(x_m) \\ \phi_2(x_m) \\ \phi_3(x_m) \\ \vdots \\ \phi_p(x_m) \end{bmatrix} \qquad \phi_n \begin{bmatrix} \phi_1(x_n) \\ \phi_2(x_n) \\ \vdots \\ \phi_p(x_n) \end{bmatrix}$$

$$\boxed{\phi_m^T \phi_n = k(x_m, x_n)}$$

If ϕ is identity $\Rightarrow \phi(x_m) = x_m$

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

What will be Φ?

Assume x_m & x_n $\mathbb{R}^2$

$x_m = [x_{m1}, x_{m2}]$

$$\begin{bmatrix} 1 \\ x_1^2 \\ x_2^2 \\ \sqrt{2}x_1 \\ \sqrt{2}x_2 \\ \sqrt{2}x_1x_2 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ x_{m1}^2 \\ x_{m2}^2 \\ \sqrt{2}x_{m1} \\ \sqrt{2}x_{m2} \\ \sqrt{2}x_{m1}x_{m2} \end{bmatrix}$$

x^* – Test instance

$$w^T x^* + b \geq 0 \qquad y^* = +1$$
$$< 0 \qquad y^* = -1$$

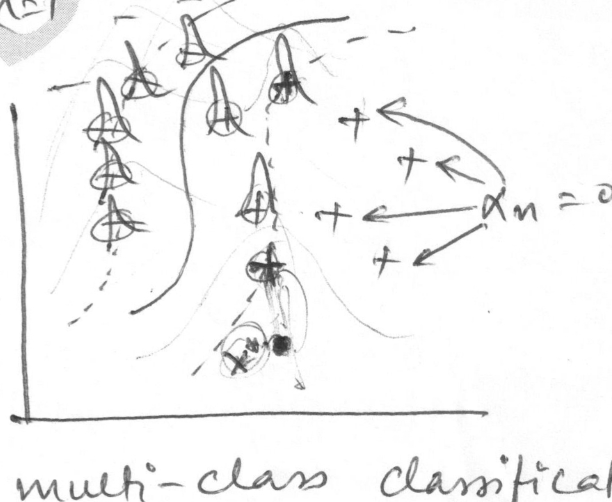
$$y^* = \text{sign}(w^T x^* + b)$$

$$w = \sum_{n=1}^{N} \alpha_n y_n x_n$$

$$y^* = \text{sign}\left[\sum_{n=1}^{N} \alpha_n y_n x_n^T x^* + b\right]$$

for Kernel SVM

$$y^* = \text{sign}\left[\left(\sum_{n=1}^{N} \alpha_n y_n k(x_n, x^*)\right) + b\right]$$



Extending SVM to multi-class classification.

let $y = \{0, 1, 2\}$
Solve three binary class problems:

0 vs 1,2 | 1 vs 0,2 | 2 vs 0,1

One vs. Rest