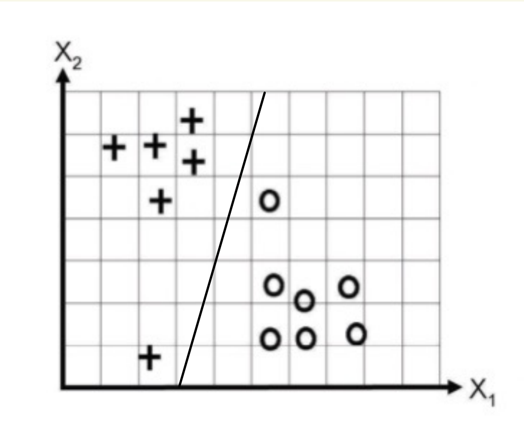
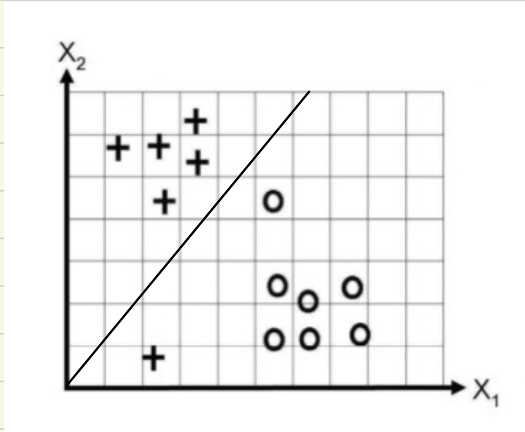
1 (1)



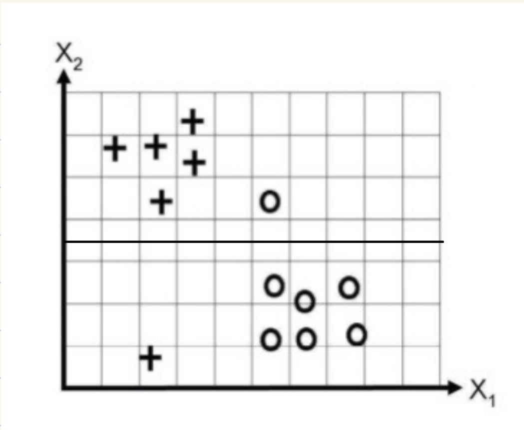
0 error

(2)



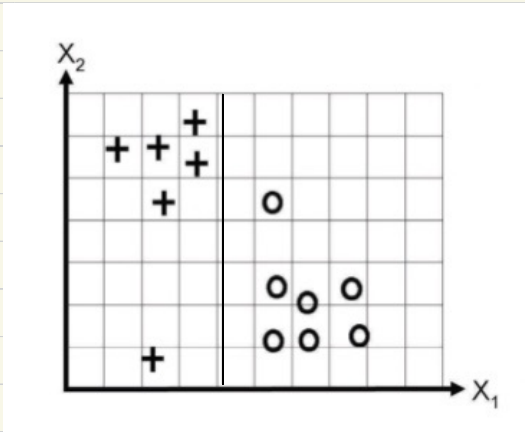
1 error

(3)



2 errors.

(4)



0 error

2.

(1) $\phi(x_1) = [1, -1, 1]^T$, $\phi(x_2) = [1, 2, 4]^T$

$\Delta\phi = [0, 3, 3]$ $\Rightarrow w = [1, 0, 0]^T \times [0, 3, 3]^T = [0, -3, 3]^T$.

(2) $\text{proj}(\phi(x_1)) = \dfrac{\phi(x_1) w}{\|w\|^2} w = \dfrac{[1, -1, 1]^T \cdot [0, -3, 3]^T}{18} \cdot [0, -3, 3]^T = [0, 1, -1]^T$

$\text{proj}(\phi(x_2)) = \dfrac{\phi(x_2) w}{\|w\|^2} w = \dfrac{[1, 2, 4]^T \cdot [0, -3, 3]^T}{18} \cdot [0, -3, 3]^T = [0, 2, -2]^T$

$y_1 = -1$, $y_2 = 1$ $\Rightarrow$ $\text{margin} = \|\text{proj}(\phi(x_2)) + \text{proj}(\phi(x_1))\| \cdot \frac{1}{2} = \frac{\sqrt{18}}{2}$

(3) $\text{margin} = \dfrac{1}{\|w\|}$ $\Rightarrow \|w\| = \dfrac{2}{\sqrt{18}}$ $\Rightarrow \sqrt{18}\lambda = \dfrac{2}{\sqrt{18}}$ $\Rightarrow \lambda = \dfrac{1}{9}$

$\Rightarrow w = [0, -\frac{1}{3}, \frac{1}{3}]^T$

3.

(1) No, the set is obviously non-linear. We can use kernels to map the datas into higher dimension to find a dicision boundary.

(2) The definition of RBF considering two vars x, x' is:

$$K(x, x') = \exp\left(-\frac{1}{2\sigma^2} \|x - x'\|^2\right)$$

It can map original data into higher feature space — where non-linear data in original feature space can become linear in the new feature space.

Thus we can use SVM to deal with it.

Its advantage is it can deal with non-linear sets like the given data set.

4. (1) $\gamma = \frac{1}{\|w\|} = \frac{1}{\sqrt{w^T w}}$

$$w^T w = \left(\sum_{i=1}^N \alpha_i y_i \phi(x_i)\right)^T \left(\sum_{j=1}^N \alpha_j y_j \phi(x_j)\right)$$

$$= \sum_{i=1}^N \sum_{j=1}^N \alpha_i y_i \alpha_j y_j \phi(x_i)^T \phi(x_j)$$

$$= \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

ϕ effects K, then effects γ.

$$\Rightarrow \gamma = \frac{1}{\sqrt{\sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j)}}$$

(2) $$\frac{1}{\gamma^2} = \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

$$= \sum_{i=1}^N \alpha_i \alpha_i y_i y_i K(x_i, x_i) + \sum_{i=1}^N \sum_{i \neq j} \alpha_i \alpha_j y_i y_j K(x_i, x_j)$$

Since $w = \sum_{i=1}^N a_i y_i \phi(x_i)$

$\Rightarrow$ For $\forall i \neq j$, $j \notin S$ $\Rightarrow \alpha_j = 0$.

$$\Rightarrow \frac{1}{\gamma^2} = \sum_{i=1}^N \alpha_i \alpha_i y_i y_i K(x_i, x_i)$$

Since $y_i = \pm 1 \Rightarrow y_i^2 = 1$

$$\Rightarrow \frac{1}{\gamma^2} = \sum_{i=1}^N \alpha_i \alpha_i K(x_i, x_i)$$

For $i \in S$, $y_i(w^T \phi(x_i) + b) = 1$

$w = \sum_{i=1}^N \alpha_i y_i \phi(x_i)$

$$\Rightarrow y_i\left(\sum_{j=1}^N a_j y_j K(x_j, x_i) + b\right) = 1$$

$$\Rightarrow \alpha_i = \frac{1 - y_i b - \sum_{j \neq i} a_j y_j K(x_j, x_i)}{K(x_i, x_i)}$$

Take it into $\sum_{i=1}^N \alpha_i^2 K(x_i, x_i)$

$$\Rightarrow \text{Original} = \sum_{i=1}^N \left(\frac{1 - y_i b - \sum_{j \neq i} a_j y_j K(x_j, x_i)}{K(x_i, x_i)}\right) \alpha_i K(x_i, x_i)$$

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

$\Rightarrow \text{Original} = \sum_{i=1}^N \alpha_i \quad \Rightarrow \frac{1}{\gamma^2} = \sum_{i=1}^N \alpha_i$