SVM

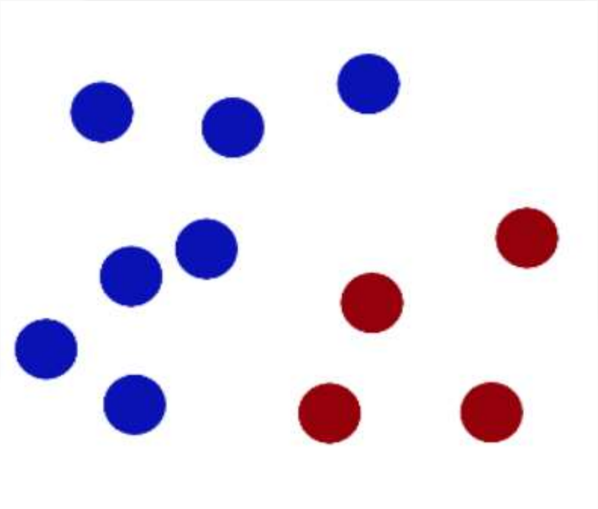
Soft-margin SVM

SVM

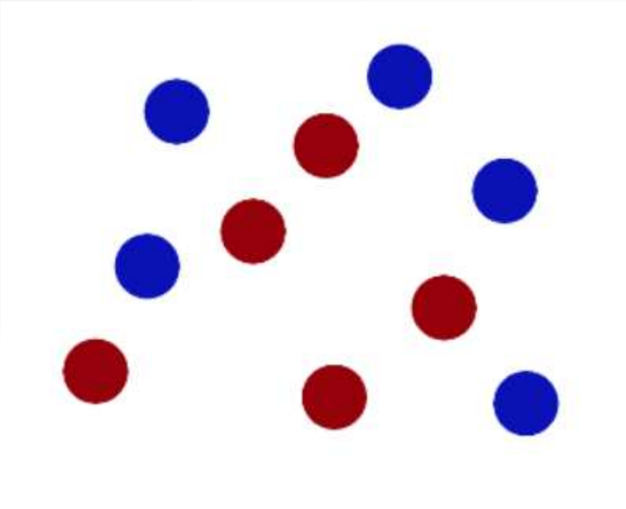
功能

主要可以解決：

線性可分類問題



非線性可分類問題



SVM

Optimization

- 目的是為了找到最好的Hyperplane(最大margin)

SVM

Optimization

hyperplane：$$w^T \cdot x + b = 0$$

- 將式子看成向量形式。

例子：$2x + 3y + z + 1 = 0$

$w = (2, 3, 1)$, $x = (x, y, z)$, $b = 1$

※法向量

SVM

Optimization

二維空間點（x, y）到直線$ax+by+c=0$的距離公式是：

$$\frac{|Ax+By+C|}{\sqrt{A^2+B^2}}$$

延伸

n維空間點（x1, x2,… ,xn）到 $w^T \cdot x + b = 0$ 的距離公式是：

$$\frac{|w^T x + b|}{||w||}$$

$$||w|| = \sqrt{\{w_1^2, w_2^2, w_3^2, \ldots, w_n^2\}}$$

SVM

Optimization

$$d = \frac{|w^T x + b|}{||w||}$$

根據Support vector的定義， Support vector到超平面的距離為margin，且其他點到超平面的距離大於margin。

$$\Rightarrow \begin{cases} \dfrac{w^T x + b}{||w||d} \geq 1 & y = 1 \\ \dfrac{w^T x + b}{||w||d} \leq -1 & y = -1 \end{cases}$$

SVM

Optimization

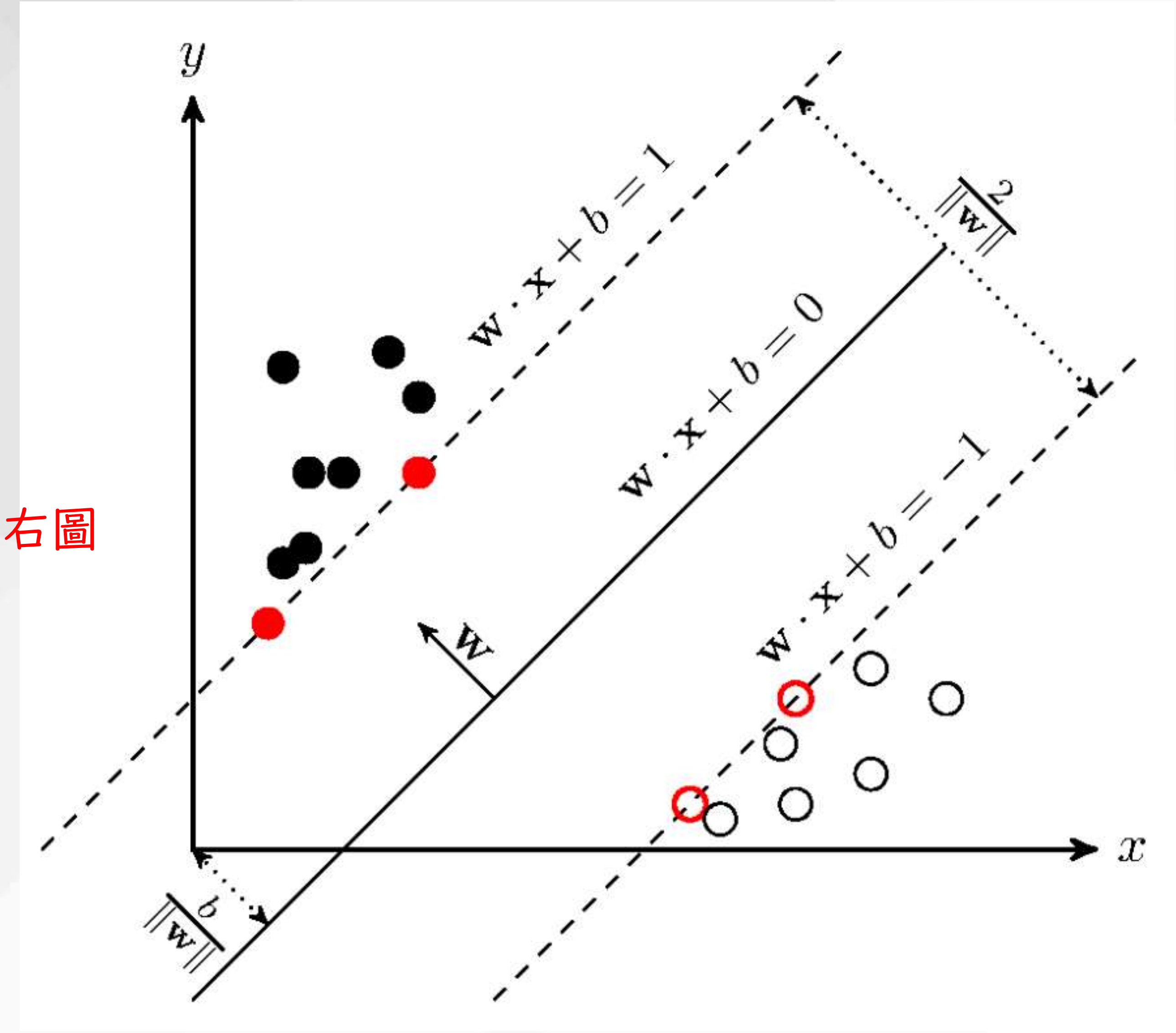
$$\begin{cases} w^T x + b \geq 1 & y = 1 \\ w^T x + b \leq -1 & y = -1 \end{cases}$$

↓ 簡化

※藉著這個公式並帶入點可以得到右圖

$$y(w^T x + b) \geq 1$$

將Data set裡的點帶入公式，
≥ 1的分成一群，≤ 1的分成一群



SVM

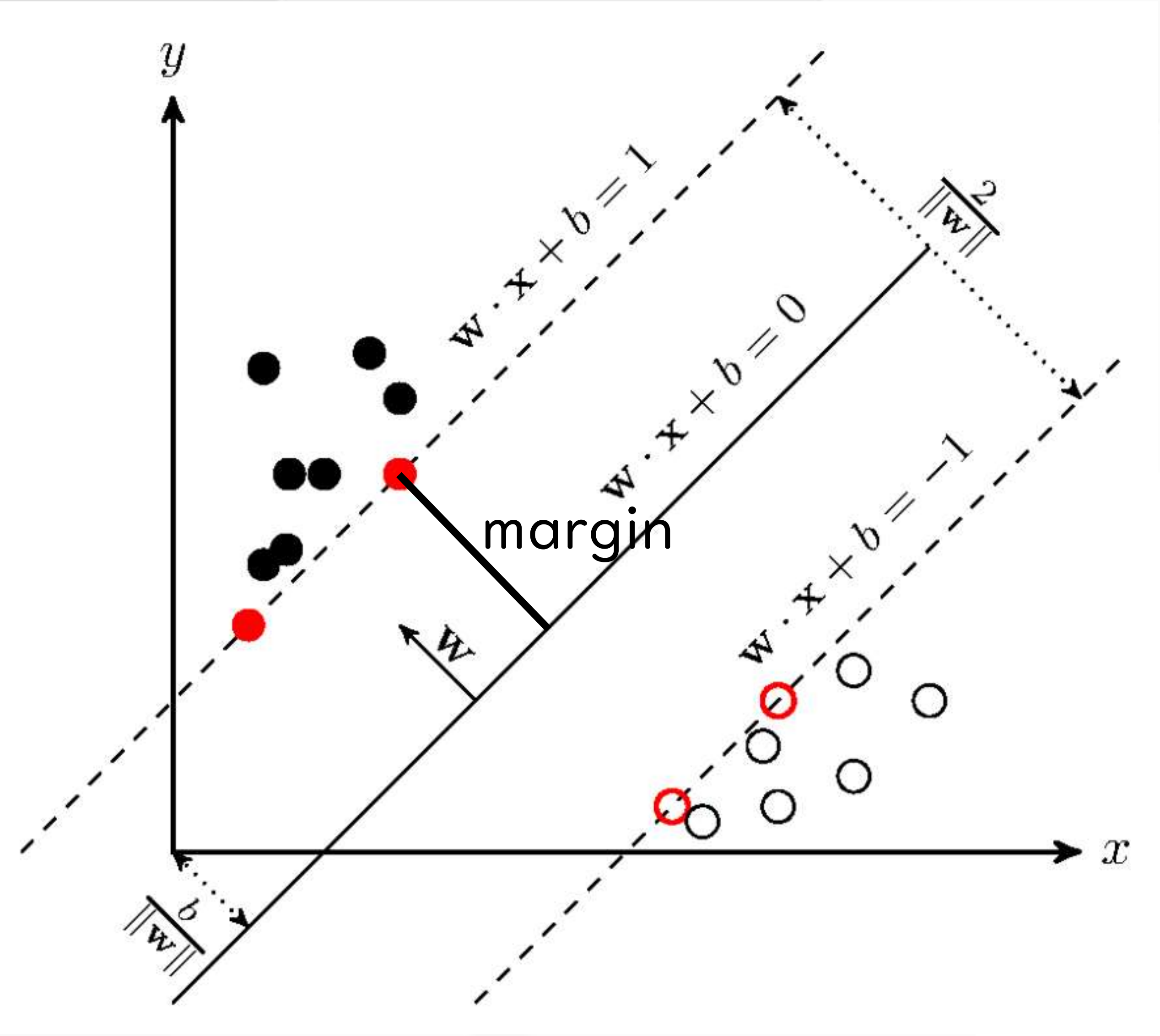
Optimization

Recall：目的是要找最大margin，margin是距離

margin = support vector到hyperplane的距離

$$= \frac{|w^T x + b|}{||w||}$$

$$= \frac{y(w^T x + b)}{||w||}$$



SVM

Optimization

得到margin = $\frac{y(w^Tx+b)}{||w||}$

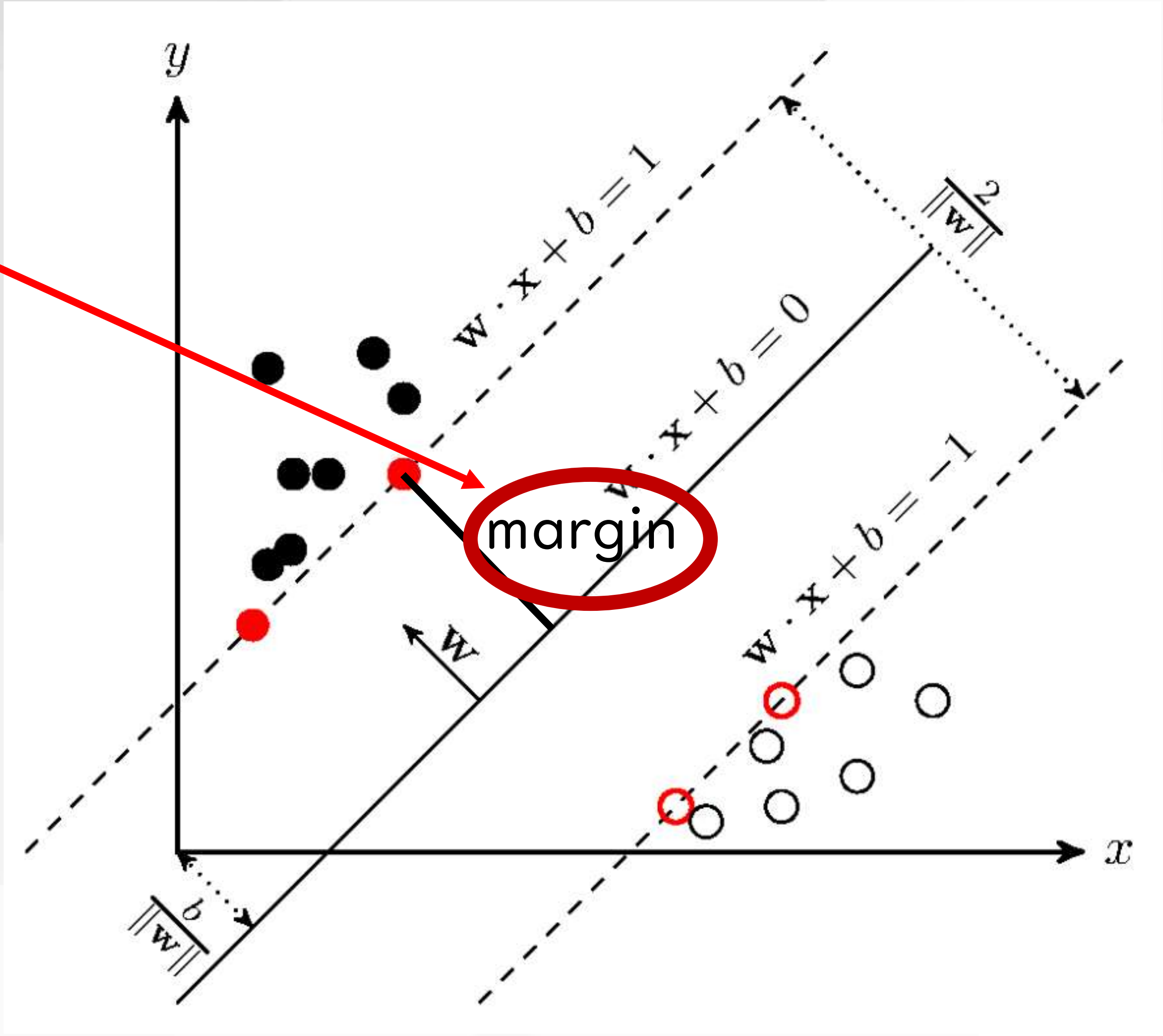
而SVM求最大margin

$$\Rightarrow \max 2 * \frac{y(w^Tx+b)}{||w||}$$

※方便推導用

※一長串東西 = 1

$$\Rightarrow \max \frac{2}{||w||}$$



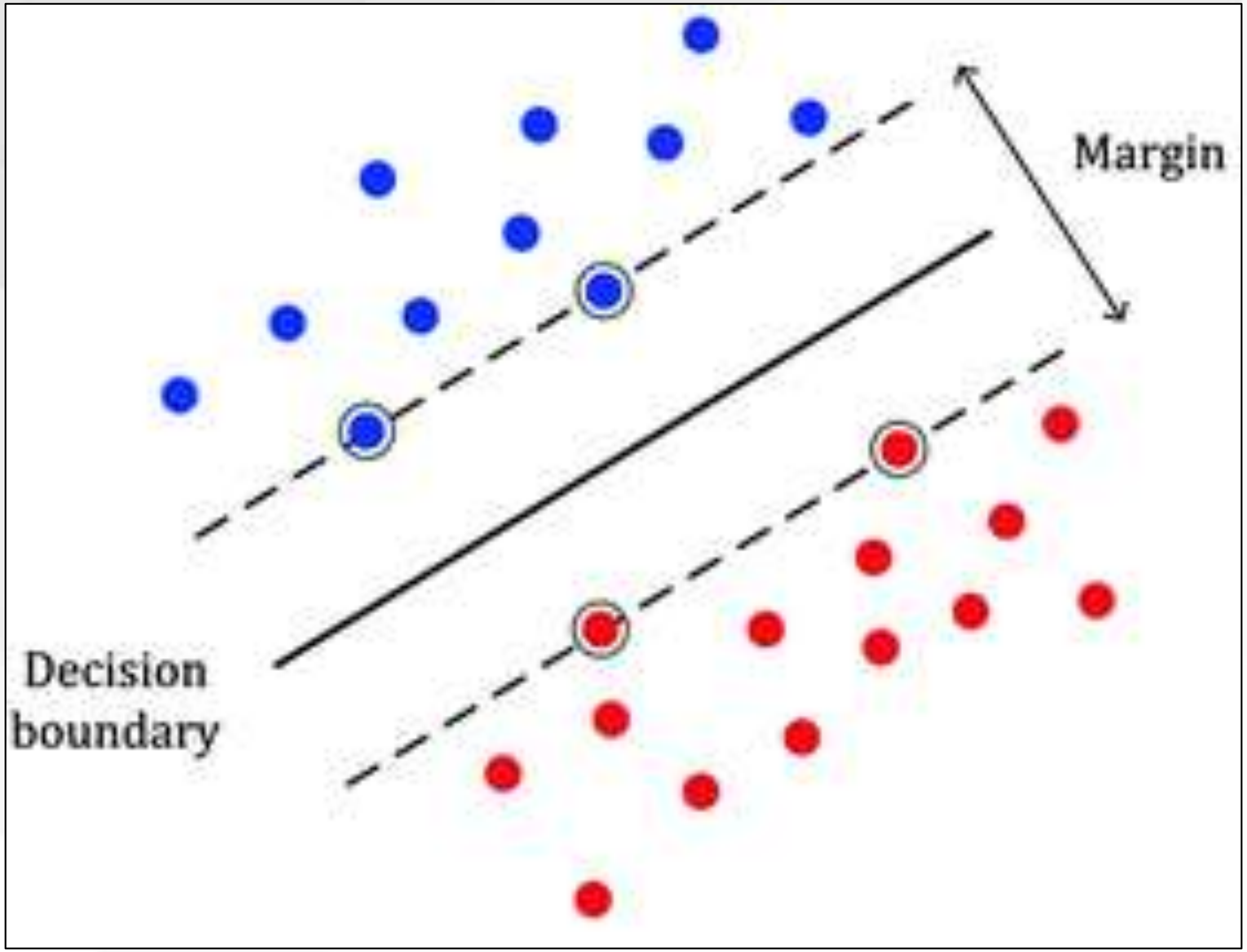
SVM

Optimization

最後得到我們知道我們要求的是：

$$\max \frac{2}{||w||} \quad \Rightarrow \quad \min \frac{1}{2}||w|| \quad s.t. \quad y_i \ (w^T x_i + b) \ \geq 1$$

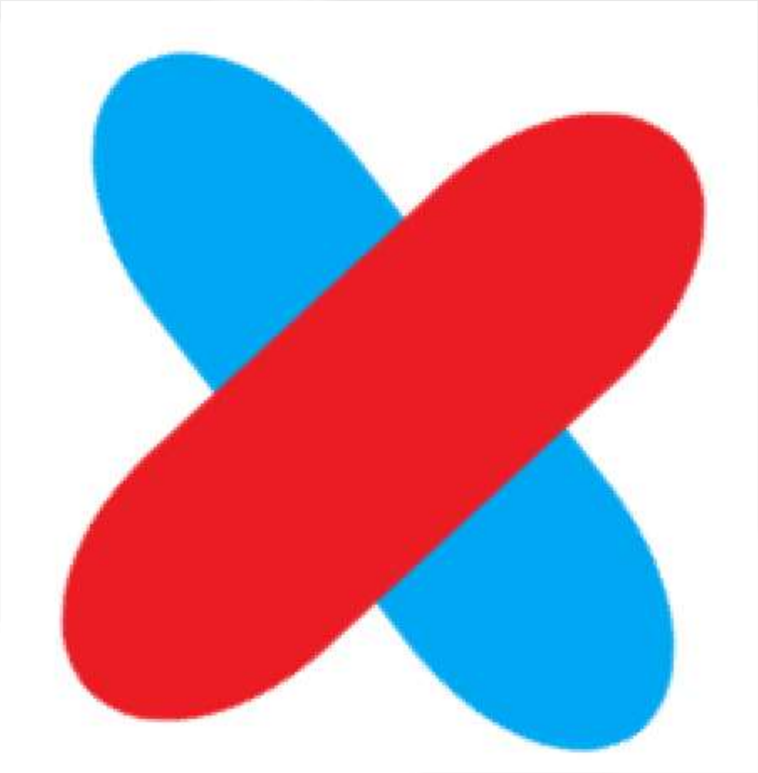
Hard-margin

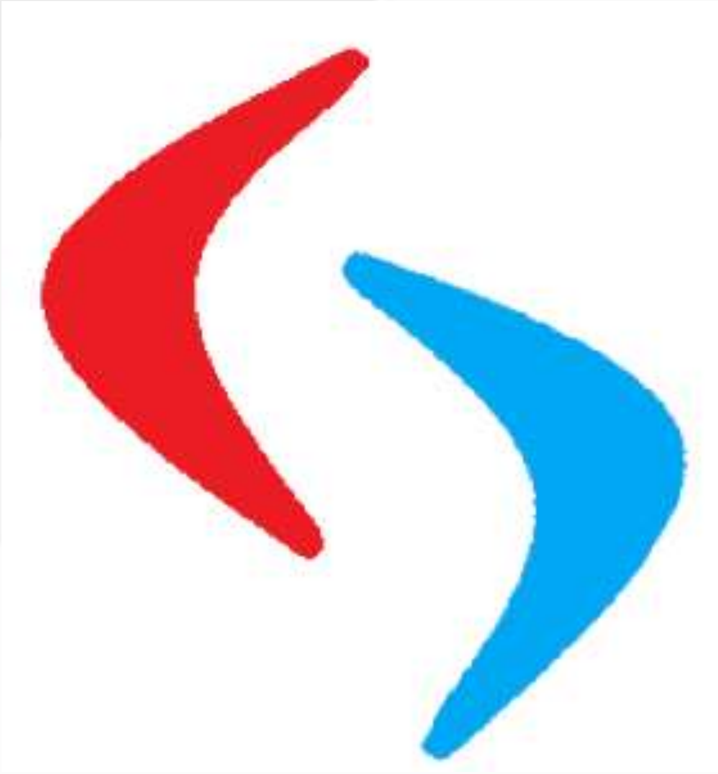


Hard-margin

Problems

- 找不到完美的hyperplane

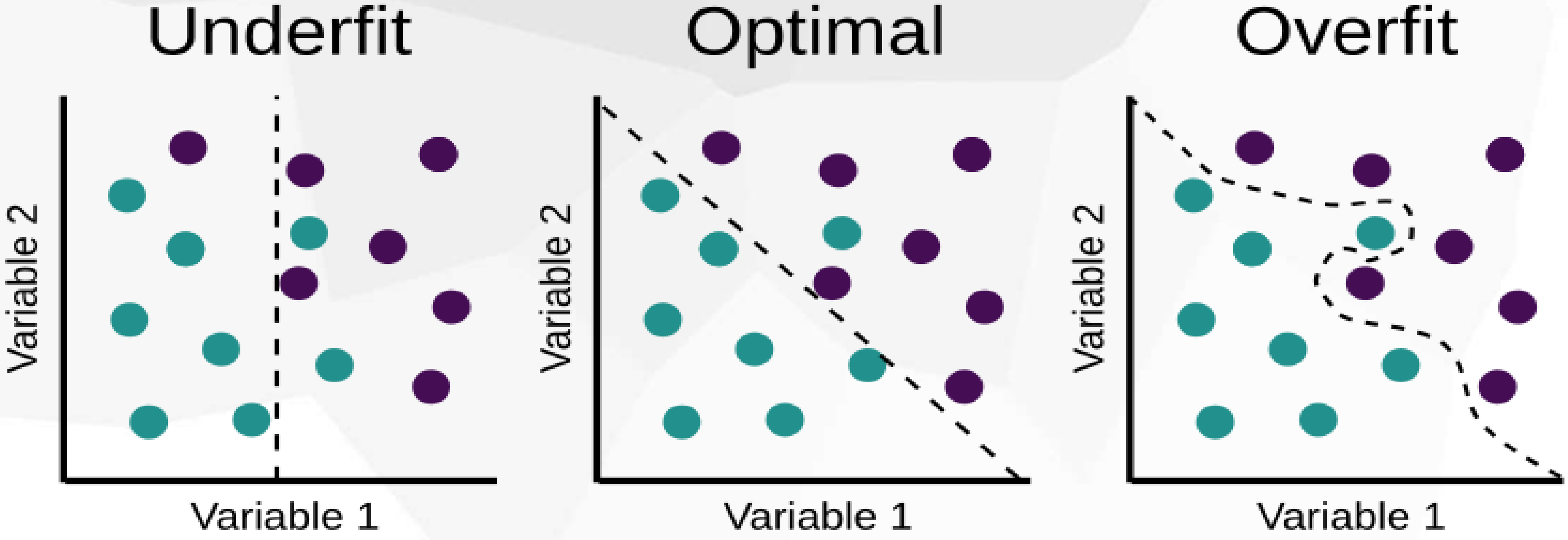




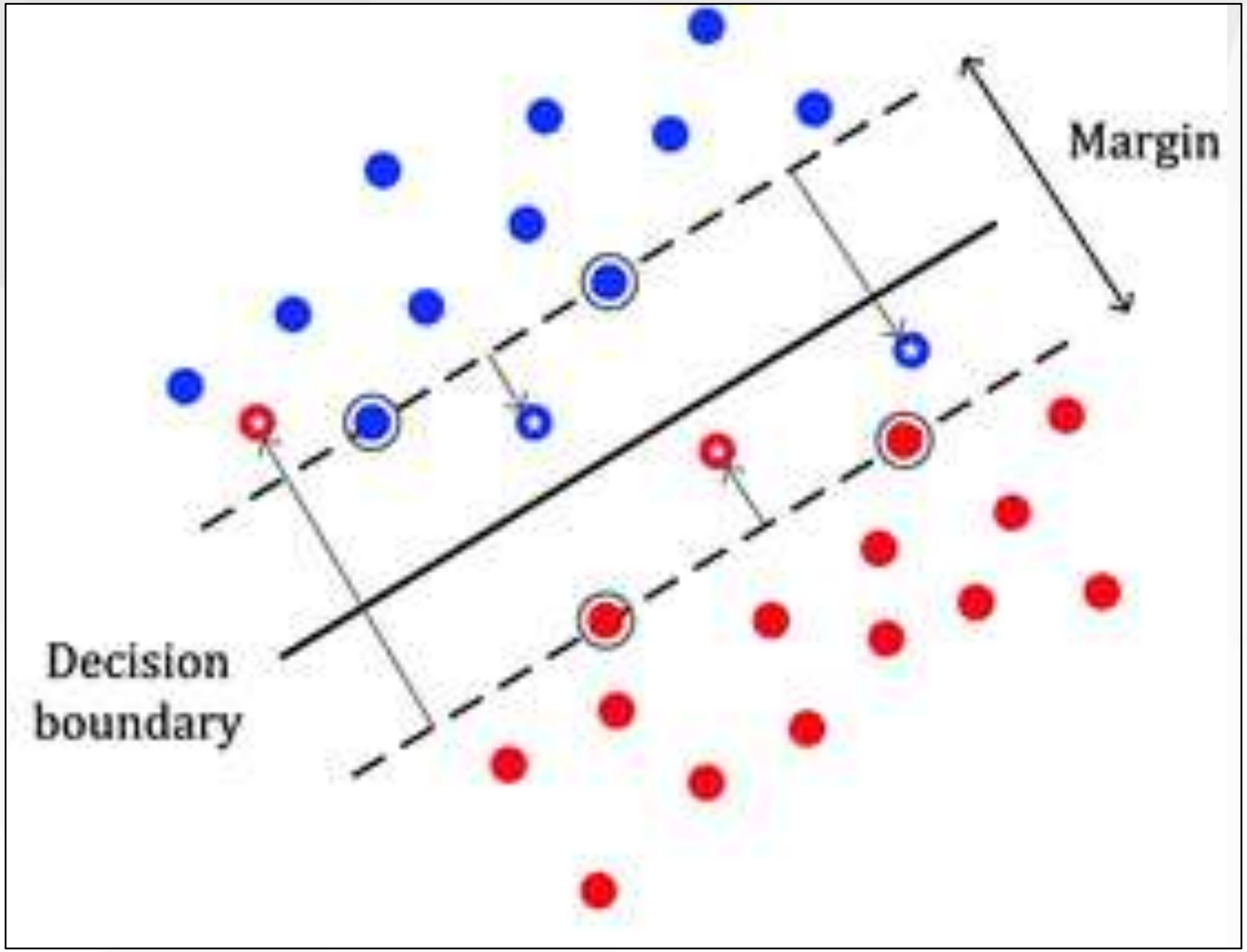
Hard-margin

Problems

- Overfitting



Soft-margin



Soft-margin

difference

加入了懲罰參數"c" & 容忍誤差"ξ" (ksi)

$$\min_{w,\xi_i} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i$$

$$y_i(w^T \boldsymbol{x}_i + b) \geq 1 - \xi_i$$

Soft-margin

difference

ξ (ksi)

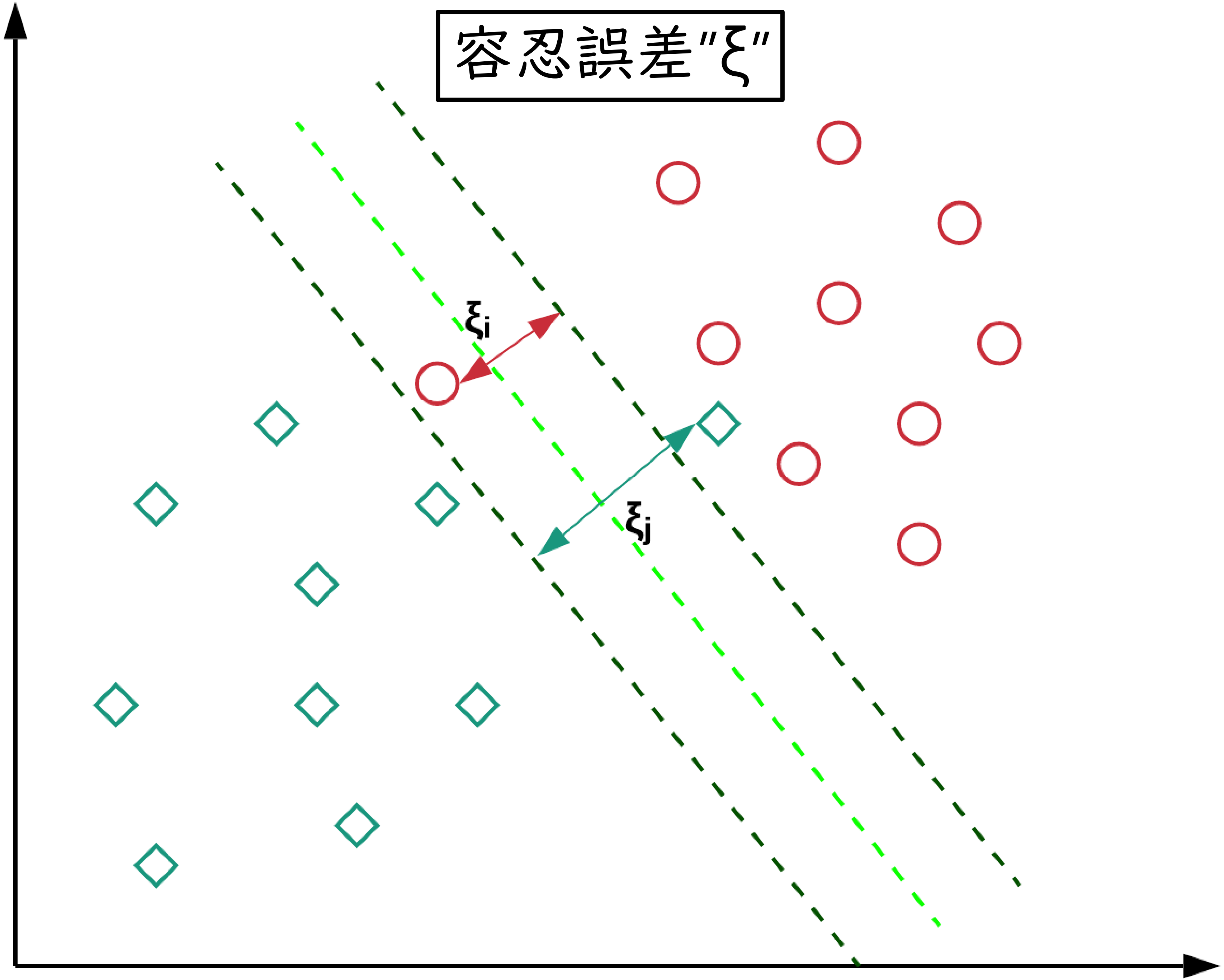
容忍誤差"ξ"：錯誤的Data與邊界的距離

懲罰參數"c"：負責控制ξ總和的影響程度

$$\min_{w,\xi_i} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i$$

$$y_i(w^T \mathbf{x}_i + b) \geq 1 - \xi_i$$

容忍誤差"ξ"



$\boxed{\text{懲罰參數"c"}}$

$$\min_{w,\xi_i} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i$$

$$y_i(w^T \boldsymbol{x}_i + b) \geq 1 - \xi_i$$

C increase $\Rightarrow \sum_{i=1}^{n} \xi_i$ decrease $\Rightarrow \xi_i$ decrease

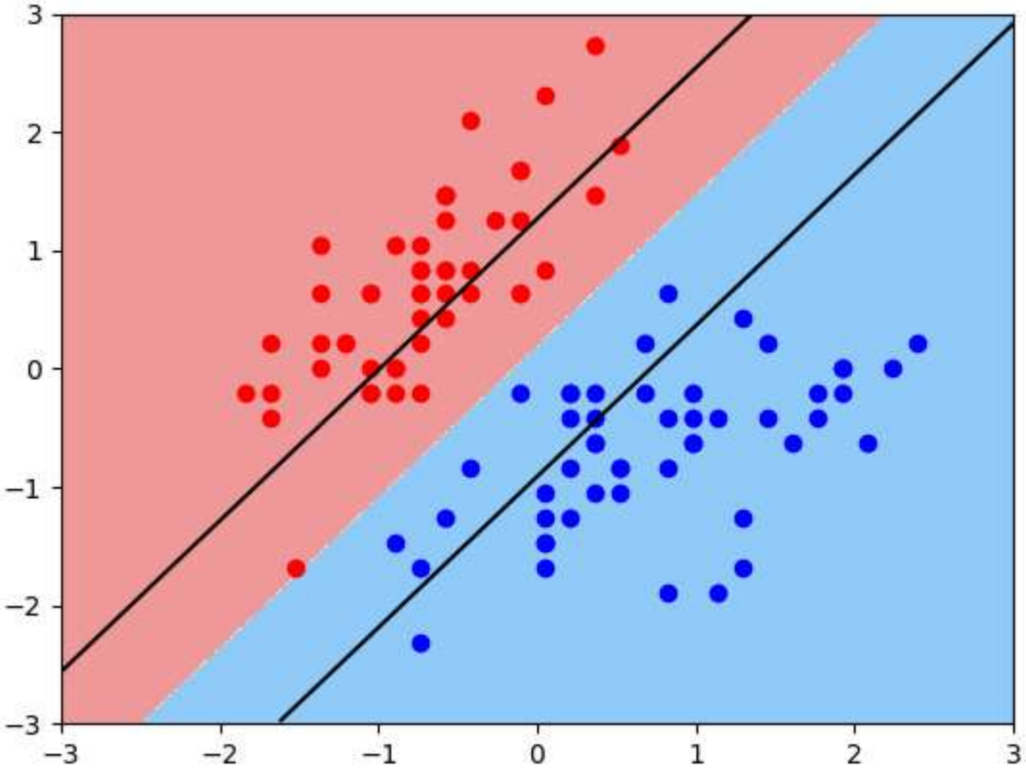
$\Rightarrow$ Margin decrease

※ c越大越嚴格，越小越鬆散

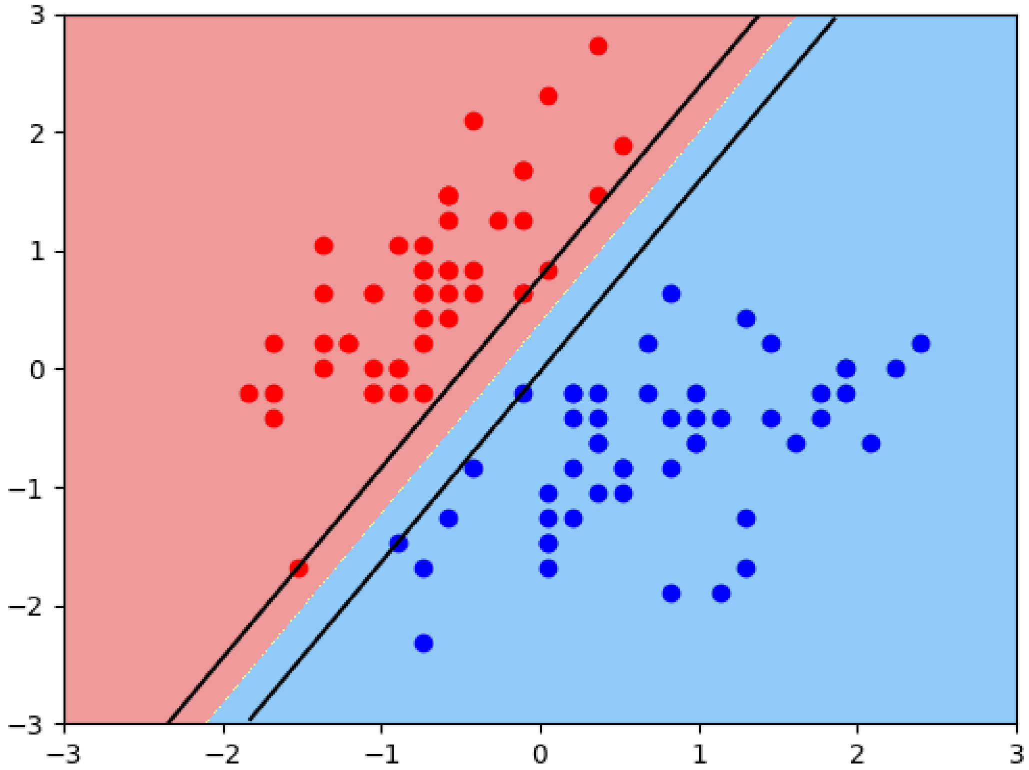
懲罰參數"c"

$$\min_{w,\xi_i} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i$$

C = 1



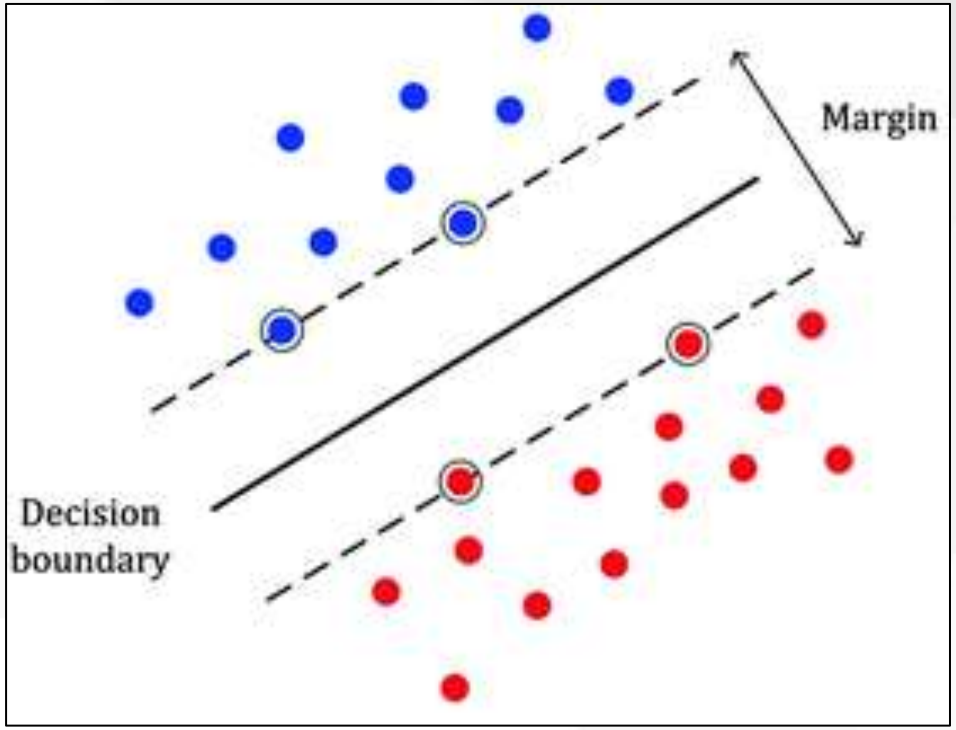
C = 1e9



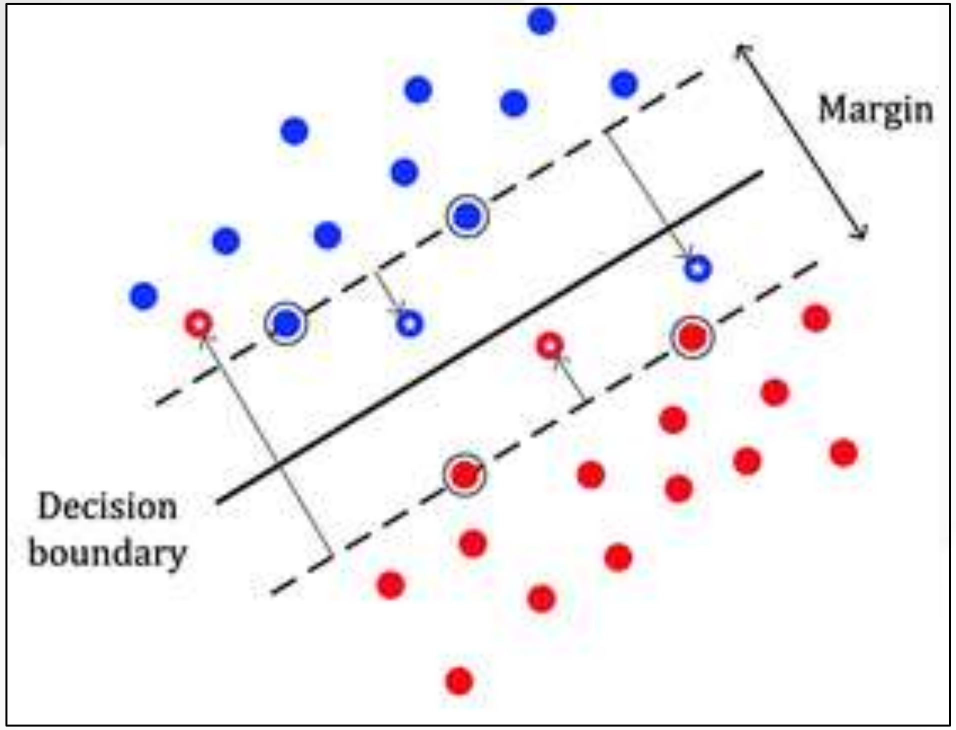
SVM

比較-圖型

Hard-margin



Soft-margin



SVM

比較-數學

Hard-margin

$$\min_{w} \frac{1}{2} w^T w$$

Soft-margin

$$\min_{w,\xi_i} \frac{1}{2} w^T w + C \sum_{i=1}^{n} \xi_i$$