→ SOM has 3 stages, ① Comptetion
② Collaboration
③ update weights

Comptetion

Find the most similar unit

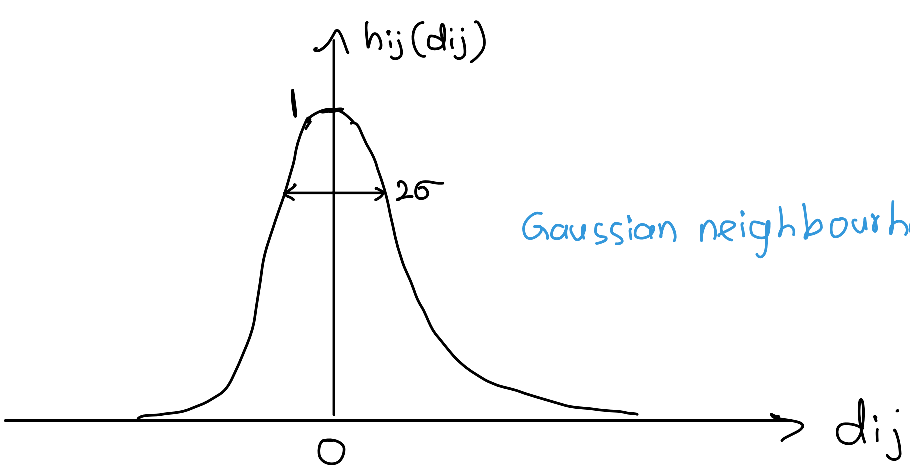
$$i(x) = \arg\max_{j} \| x - w_j \|_2 \qquad j = 1, 2, 3 \ldots n \; ; \; m \neq \text{units}$$

Collaboration

Use the latteral distance d_{ij} between the winner unit i and unit j

$$h_{i,j}(d_{ij}) = \exp\left(\frac{-d_{ij}^2}{2\sigma^2}\right)$$

Gaussian neighbourhood



Gaussian neighbourhood

$$\sigma(n) = \sigma_0 \exp\left(\frac{-n}{T}\right)$$

n = number of iterations

T = constant

Weight updates

$$w_j(n+1) = w_j(n) + \Delta w_j$$

n = iteration

constant out in adjustment

$$\Delta w_j = \underbrace{\eta \, y_i \, x}_{\text{hebb's rule}} - \underbrace{g(y_i)\, w_j}_{\text{forgetting rule}}$$

forgetting rule → bring a balance between competition and collaboration

$$g(y_i) = \eta \, y_i = \eta \, \underline{h_{ij}(x)}$$ → neighbourhood function

$$w_j(n+1) = w_j(n) + \eta(n) \, h_{ij(x)}(n) \, [x - w_j(n)]$$

$$\eta(n) = \eta \cdot \exp\left(\frac{-n}{T_2}\right)$$

constant

$\eta_0 = 0.1$

$T_2 = 1000$

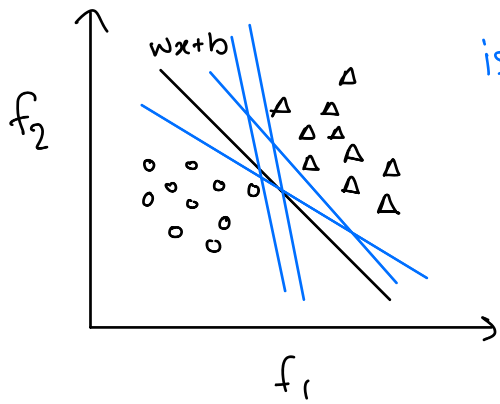
$$T_1 = \frac{1000}{\log(\underline{\sigma_0})}$$ → big enough

Converge: Many iterations
(eg: several thousand times the number of units)

stop: No noticable change
No big change in feature maps

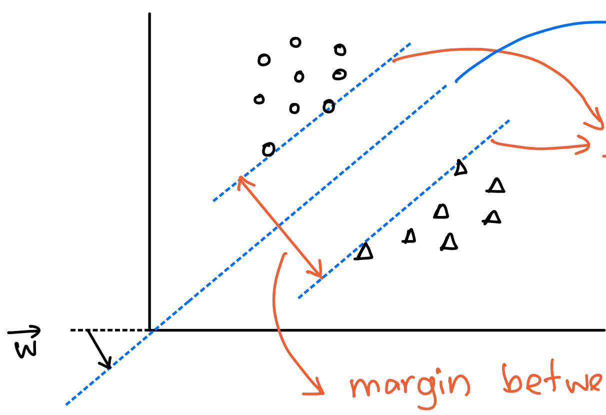
problems: Convergence take long time
Get variable results

Classification: Intelligence is to distinguish things (recognize as different)



is there an optimal line ?

Support Vector Machines (SVMs)



wx+b, line (2D), hyperplane (nD)

Support vectors

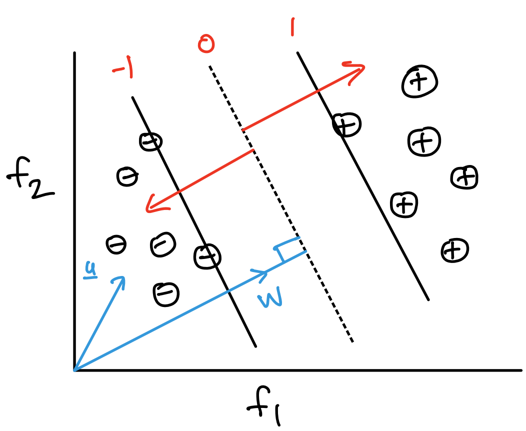
$$wx+b>0$$

$$wx+b<0$$

margin between 2 classes

$w \perp$ middle line

Assumption: classes $\in \{\oplus, \ominus\}$



$$wx_{\oplus}+b \geq 1$$

$$wx_{\ominus}+b \leq -1$$

$w \cdot x_{\oplus} + b \geq 1$

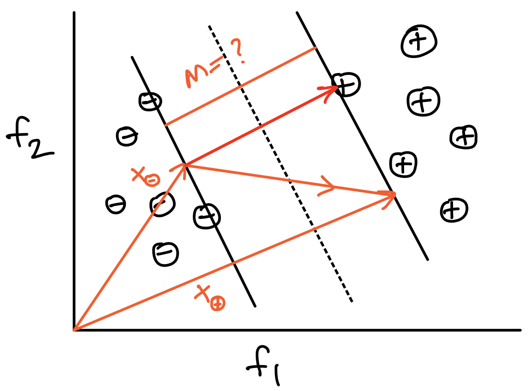
$w \cdot x_{\ominus} + b \leq -1$

Introduce a dummy variable for class membership

$$\left.\begin{array}{l} y_i = +1 \quad \text{for } \oplus \\ y_i = -1 \quad \text{for } \ominus \end{array}\right\} \begin{array}{l} y_i(w x_{\oplus} + b) \geq 1 \\ y_i(w x_{\ominus} + b) \geq 1 \end{array}$$

$$y_i(w \cdot x + b) - 1 \geq 0$$

Best classification ⟶ largest margin



$(x_{\oplus} - x_{\ominus}) \cdot \frac{w}{\|w\|}$

$$(x_{\oplus} - x_{\ominus}) \frac{w}{\|w\|} = \frac{w \cdot x_{\oplus} - w x_{\ominus}}{\|w\|} = \frac{(1-b) + (1+b)}{\|w\|} = \frac{2}{\|w\|}$$

Maximize $\frac{2}{\|w\|}$

Minimize $\|w\|$ s.t $y_i(w \cdot x_i + b) - 1 \geq 0$

Lagrange Multipliers :

$$L = \frac{1}{2}\|w\|^2 - \sum \alpha_i \left[y(w \cdot x_i + b) - 1 \right]$$

α_i - Lagrange multiplier

line/plane

$$\frac{\partial L}{\partial w} = \|w\| - \sum \alpha_i y_i x_i = 0$$

$$\|W\| = \sum \alpha_i y_i x_i$$

$$\frac{\partial L}{\partial b} = -\sum \alpha_i y_i = 0 \qquad \sum \alpha_i y_i = 0$$

Substitute w in L with $\sum \alpha_i y_i x_i$, we get after simplification :

$$L = \sum \alpha_i - \frac{1}{2} \sum\sum \underbrace{\alpha_i \alpha_j y_i y_j}_{\text{scalors}} \underbrace{x_i \cdot x_j}_{\text{dot product}}$$

Minimize this via quadratic optimization

How to classify :

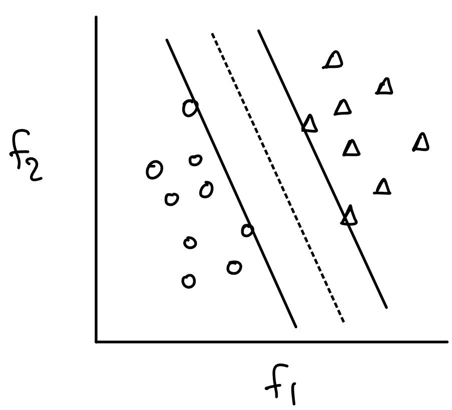
$$\sum \alpha_i x_i y_i \cdot u + b \geq 0 \longrightarrow \oplus$$

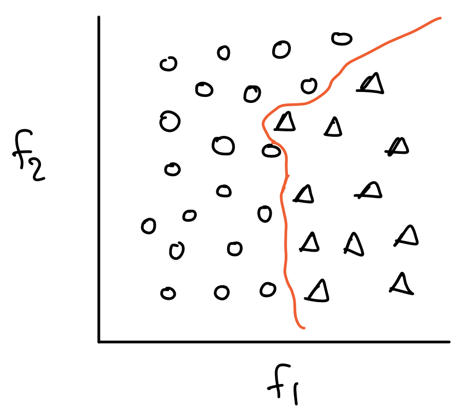
new data

otherwise $\longrightarrow \ominus$

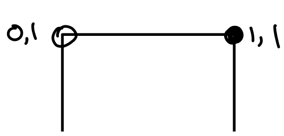
* SVM only works for binary lineary separable problems.

most real world problems

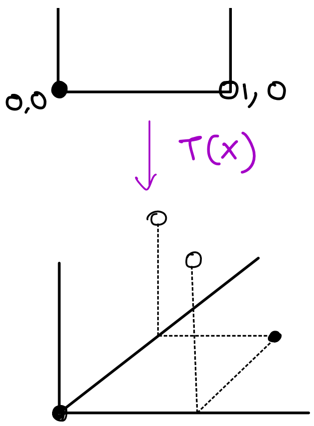




XOR is a non linear problem



Assume $T(x)$ is a transformation



that moves x to higher dimention. and making linear seperation possible. Then we have to calculate $T(x_i) * T(u)$

But this would be difficult.

If we had a function $K(x_i, x_j)$ s.t $k(x_i, x_j) = T(x_i) \cdot T(x_j)$

then we won't need T!

All we need is a kernal function K.

We don't need $T(x)$!

We just need to get $T(x_i) \cdot T(x_j)$ and not $T(x_i)$ and $T(x_j)$ individually.

The kernal trick (kernals are fundementally functions that measure similarity)

Popular kernals : $K(u,v) = (u \cdot v + 1)^n$

$$k(u,v) = \exp\left(-\frac{\|u - v\|}{\sigma}\right)$$