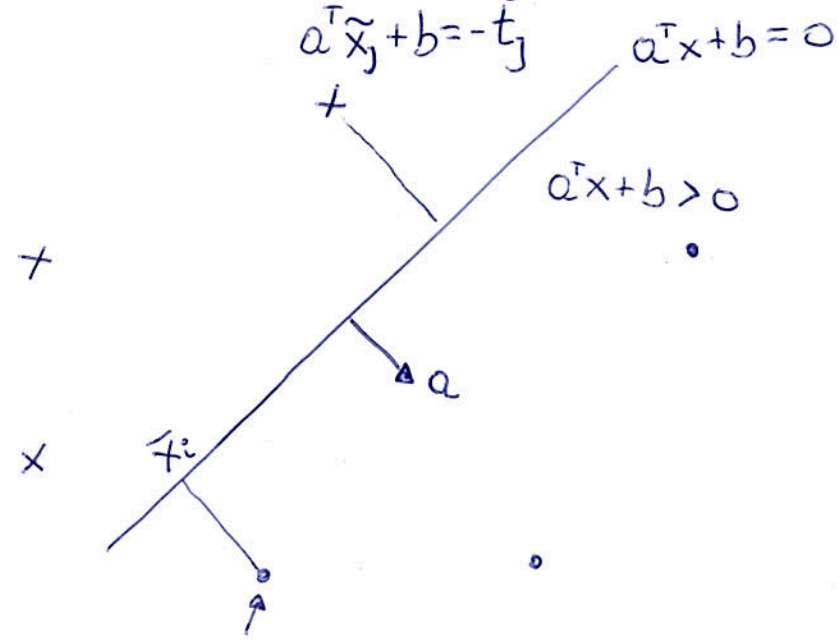
SUPPORT VECTOR MACHINE (SVM)



En base a los puntos ya clasifica[dos] se busca una recta que los separe.

Luego se usará esta recta para discriminar nuevos puntos.

$$\left.\begin{array}{l} a^Tx_i+b=t_i \\ a^T\bar{x}_i+b=0 \end{array}\right\} \quad a^T(x_i-\bar{x}_i)=t_i$$

$$\|a\|\,\|x_i-\bar{x}_i\|=t_i \Rightarrow \|x_i-\bar{x}_i\|=\frac{t_i}{\|a\|}$$

De la misma forma $\quad a^T(\tilde{x}_j-\bar{x}_j)=-\|a\|\|\tilde{x}_j-\bar{x}_j\|=-\tilde{t}_j \Rightarrow \|x_j-\tilde{x}_j\|=\frac{\tilde{t}_j}{\|a\|}$

P1)
$$\max_{a,b,t,\tilde{t}} \ \min\left\{\frac{t_1}{\|a\|},\dots,\frac{t_N}{\|a\|},\frac{\tilde{t}_1}{\|a\|},\dots,\frac{t_N}{\|a\|}\right\}$$

$$a^Tx_i+b=t_i$$
$$a^T\tilde{x}_j+b=-\tilde{t}_j$$

PARA EL CASO LINEALMENTE SEPARABLE, SE BUSCA MAXIMIZAR LA DISTANCIA DE LA RECTA A LOS PUNTOS YA CLASIFICADOS (CLASIFICACIÓN ROBUSTA)

En la próxima página transformamos (P1) en el problema equivalente (SVM-0) el cual define las SVMs en su forma estándar

(P2) $\max_{a,b,t,\tilde{t},u} \dfrac{u}{\|a\|}$

$$u \leq t_i$$
$$u \leq t_j$$

$$a^T x_i + b = t_i$$
$$a^T \tilde{x}_j + b = -\tilde{t}_j$$

(P2) equivalente a (P1) pues las desigualdades deben cumplirse con igualdad para algunos puntos, de otra forma puede agrandarse u.

(P3) $\max_{a,b,u} \dfrac{u}{\|a\|}$

$$u \leq a^T x_i + b$$
$$u \leq -a^T \tilde{x}_j - b$$

Las variables t y $\tilde{t}$ son ahora innecesarias

(P4) $\max_{a,b,u} \dfrac{u}{\|a\|}$

$$1 \leq \frac{a^T}{u} x_i + \frac{b}{u}$$
$$1 \leq -\frac{a^T}{u} \tilde{x}_j + \frac{b}{u}$$

El problema es indeterminado por su factor de escala

(P5) $\max_{a,b} \dfrac{1}{\|a\|}$

$$1 \leq a^T x_i + b$$
$$1 \leq -a^T \tilde{x}_j - b$$

Se redefinen entonces las variables $a = \frac{a}{u}$ y $b = b/u$ lo que equivale a anclar $\min(t_i) = 1$ y $\min \|x_i - \bar{x}_i\| = \dfrac{1}{\|a\|}$

$$(\text{SVM}-\emptyset) \qquad \min_{a,b} \|a\|^2$$

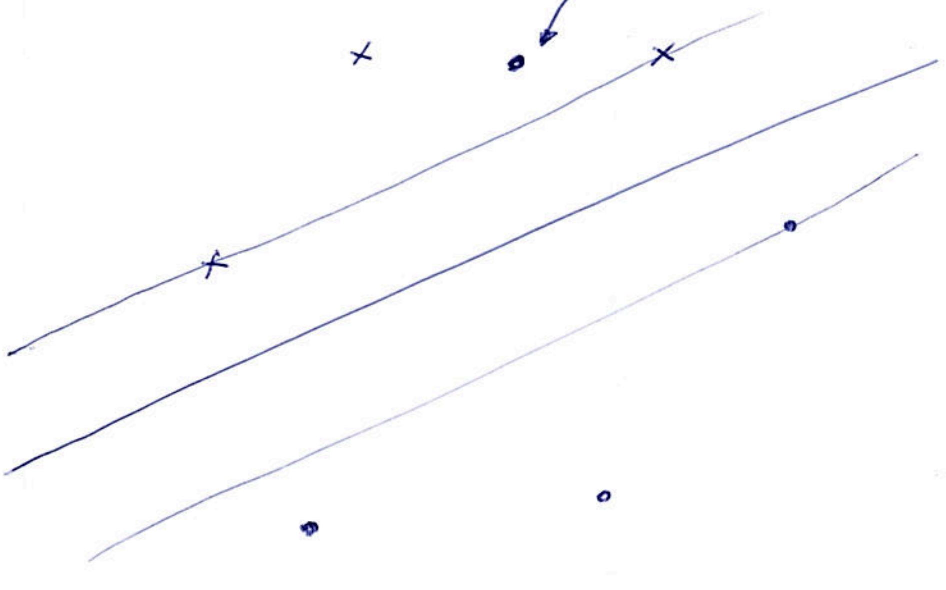
EQUIVALENTE A (P5)

$$y_i\,(a^T x_i + b) \geq 1$$

y_i: ETIQUETAS DE CLASE $\{-1, 1\}$

SOLUCIÓN

AGREGANDOLO, LA MUESTRA PASA A SER NO SEPARABLE LINEALMENTE



LA SOLUCIÓN A SVM-$\emptyset$ ESTÁ DETERMINADA POR EL CONJUNTO DE PUNTOS QUE SOSTIENEN LA FRANJA ALREDEDOR DE LA RECTA ÓPTIMA. A ESTOS PUNTOS SE LOS LLAMA SUPPORT VECTORS.

$$(\text{SVM}-\varepsilon) \qquad \min_{a,b,\epsilon} \|a\|^2 + C \sum_i \epsilon_i$$

$$y_i\,(a^T x_i + b) \geq 1 - \epsilon_i$$

$$\epsilon_i \geq 0$$

EL PROBLEMA (SVM-$\emptyset$) ES INFACTIBLE SI LOS PUNTOS NO SON LINEALMENTE SEPARABLES. EN ESTE CASO SE AGREGAN LAS VARIABLES ϵ_i QUE PERMITEN A ALGUNOS PUNTOS PASAR PARA ADENTRO DEL SLAB, O INCLUSO PARA EL OTRO LADO DE LA RECTA. ESTO SE PENALIZARÁ EN EL COSTO, YA QUE CORRESPONDE AL CASO $\epsilon_i > 0$.