Máquinas de soporte vectorial

Supongamos que tenemos dos clases de patrones linealmente separables, separados por un hiper plano (w,b) donde $w \in \mathbb{R}^n$, $b \in \mathbb{R}$ y los patrones $x \in \mathbb{R}^n$. Los patrones de clase + satisfacen $w \cdot x + b > 0$ y los de clase −, $w \cdot x + b < 0$. En este caso w es el vector ortogonal al hiper plano y b es proporcional a la distancia al origen.

Los vectores soporte son los vectores más cercanos al hiper plano de separación y se pueden elegir los parametros (w,b) de tal manera que en los vectores se tenga $w \cdot x^+ + b = 1$ y $w \cdot x^- + b = -1$ para los vectores soporte de clase + y − respectivamente. Si (w,b) cumple con lo anterior entonces $\left(\frac{w}{|w|}, \frac{b}{|w|}\right)$ mide la distancia de los patrones al hiper plano y en el caso de los vectores soporte es $\frac{1}{|w|}$. Esto implica que si se quiere encontrar el hiper plano de separación que maximice la distancia entre los vectores soporte de ambos clases, se tiene que resolver el problema:

$$w_* = \arg\max_{w} w \cdot w$$
$$y_i(w \cdot x_i + b) \geq 1$$

, donde w_* determina el hiper plano óptimo buscado

para resolver este problema es suficiente encontrar un punto estacionario de la función Lagrangeana

$$L(w,b,\alpha) = \frac{1}{2}(w \cdot w) - \sum_{i=1}^{\ell} \alpha_i [y_i((w \cdot x_i) + b) - 1]$$

en terminos de w y α y b

los datos son $((x_i, y_i))$, $x_i \in \mathbb{R}^n$ son los patrones y $y_i \in \{-1, 1\}$ son las etiquetas.

Todo lo anterior de acuerdo al Teorema de Kuhn-Tucker.

En el caso que sea necesario se busca una funcion que separe linealmente a las clases.