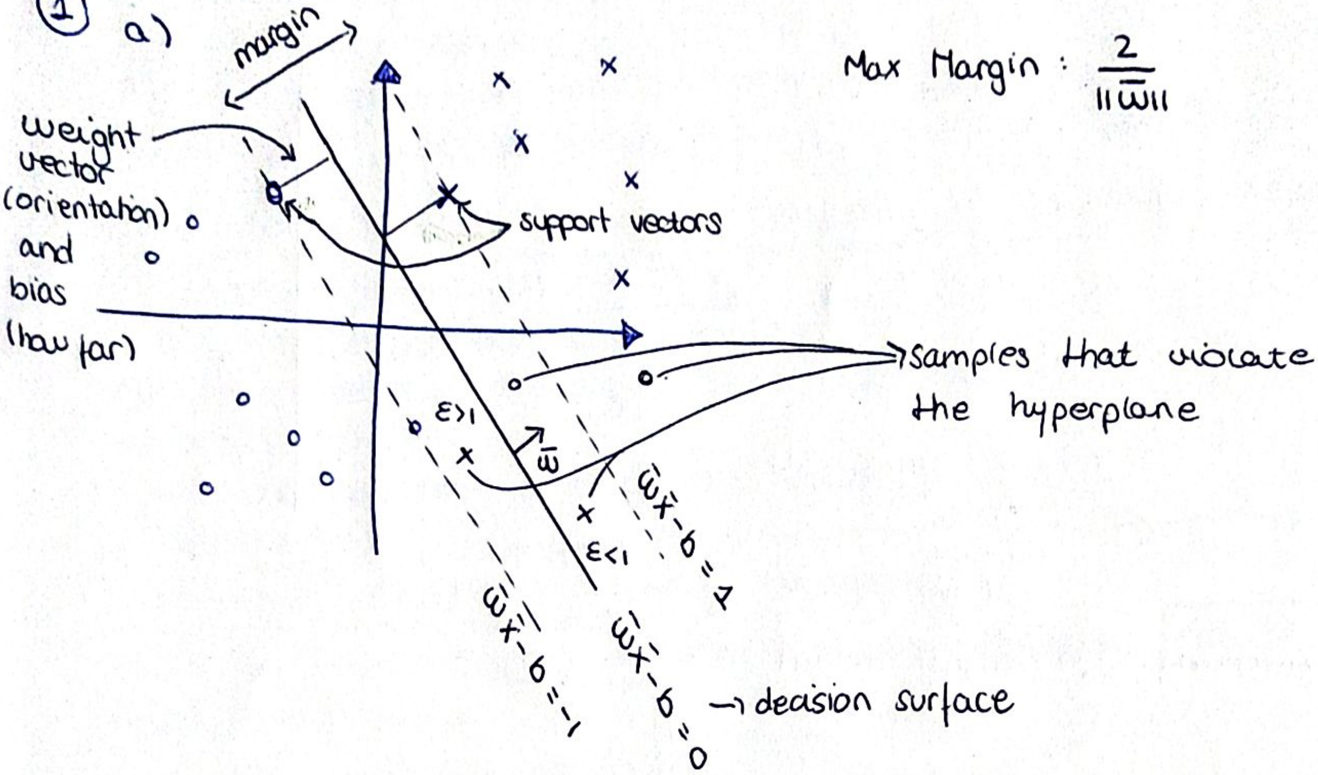
Lucía Rivas

ADVANCED COURSE IN ML

(1) a)



Max Margin : $\frac{2}{\|\bar{w}\|}$

b) $\min \frac{1}{2}\|w\|^2$ such that $(w^T x_n + b) y_n \geq 1$ is equivalent to

$\max_\alpha \left(\sum_i \alpha_i - \frac{1}{2} \sum_{ij} \alpha_i \alpha_j y_i y_j x_i^T x_j \right)$ st $\alpha_i \geq 0 \ \forall i$ and $\sum_i \alpha_i y_i = 0$

<u>Proof</u>: First we calculate the Lagrangian L :

$$L = \frac{1}{2}\|\bar{w}\|^2 - \sum_i^n \alpha_i (y_i(\bar{w}\bar{x} + b) - 1)$$

$$\frac{dL}{\partial w} = \bar{w} - \sum_i^n \alpha_i y_i x_i = 0 \implies \boxed{\bar{w} = \sum_i^n \alpha_i y_i x_i} \quad (1)$$

$$\frac{dL}{db} = -\sum_i^n \alpha_i y_i = 0 \implies \boxed{\sum_i^n \alpha_i y_i = 0} \quad (2)$$

Now
$$L = \frac{1}{2}\|\bar{w}\|^2 - \sum_i^n \alpha_i (y_i(\bar{w}\bar{x} + b) - 1) \overset{(1)}{=}$$
$$= \frac{1}{2}\left(\sum_i^n \alpha_i y_i \bar{x}_i\right)\left(\sum_j^n \alpha_j \bar{x}_j y_j\right) - \sum_i^n \alpha_i y_i \left(\sum_i^n \alpha_i y_i \bar{x}_i\right)\bar{x} + \overset{(2)}{\cancel{\sum_i^n \alpha_i y_i}} b + \sum_i^n \alpha_i$$
$$\overset{(2)}{=} \frac{1}{2}\sum_i^n \sum_j^n \alpha_j y_j \alpha_i y_i \bar{x}_i \bar{x}_j - \sum_i^n \sum_j^n \alpha_j y_j \alpha_i y_i \bar{x}_i x_j - 0 + \sum_i^n \alpha_i =$$
$$= \sum_i^n \alpha_i - \frac{1}{2}\sum_i^n \sum_j^n \alpha_i \alpha_j y_i y_j x_i x_j$$

Therefore is the same as $\max_\alpha \left(\sum_i \alpha_i - \frac{1}{2}\sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i x_j \right)$