2021-06-12

Εργασία 3: έχει ανακοινωθεί.

Μέθοδος Support Vector Classification

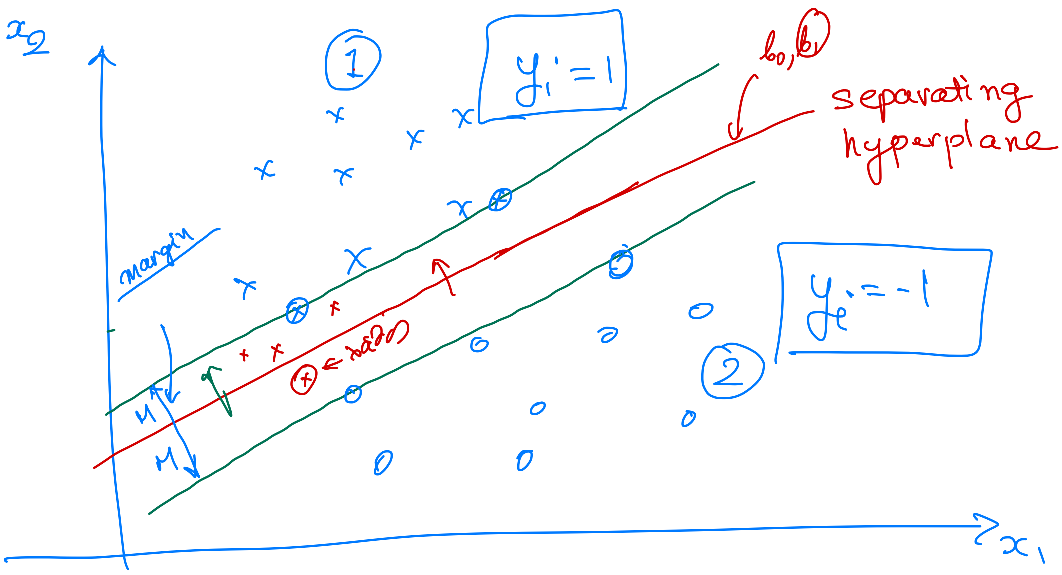
1) Separating Hyperplanes (κεφ. 4.5)
2) Κεφ 12 (SV Classifiers)

Γενίκευση: SV Machines → project

Classification με 2 κατηγορίες

Y: outcome, $X=(X_1, X_2)$: predictors

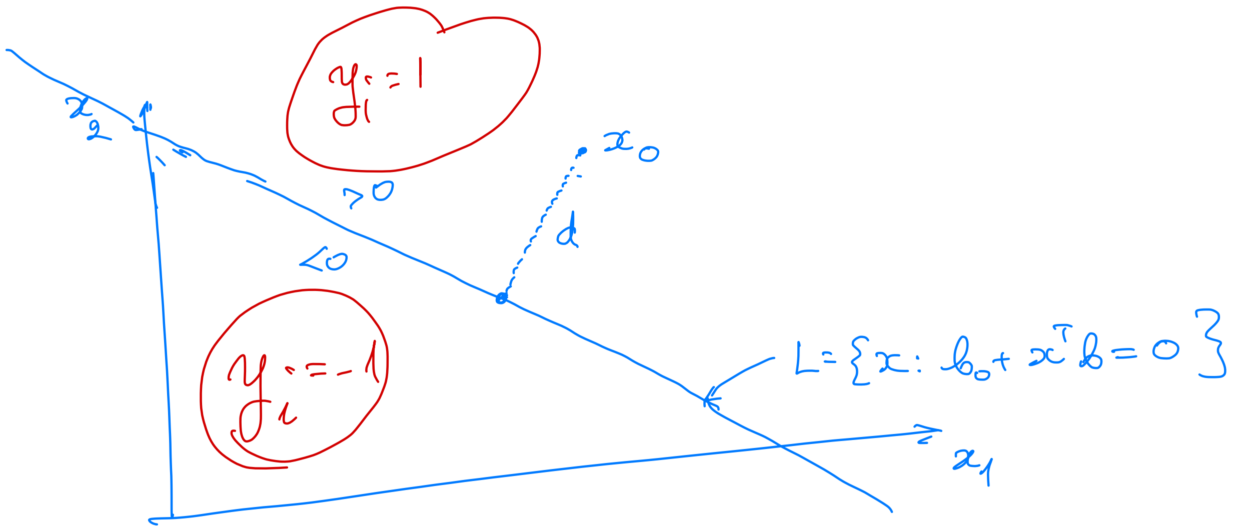
Περίπτωση 1: Training set: οι δύο κλάσεις είναι πλήρως γραμμικά διαχωρίσιμες



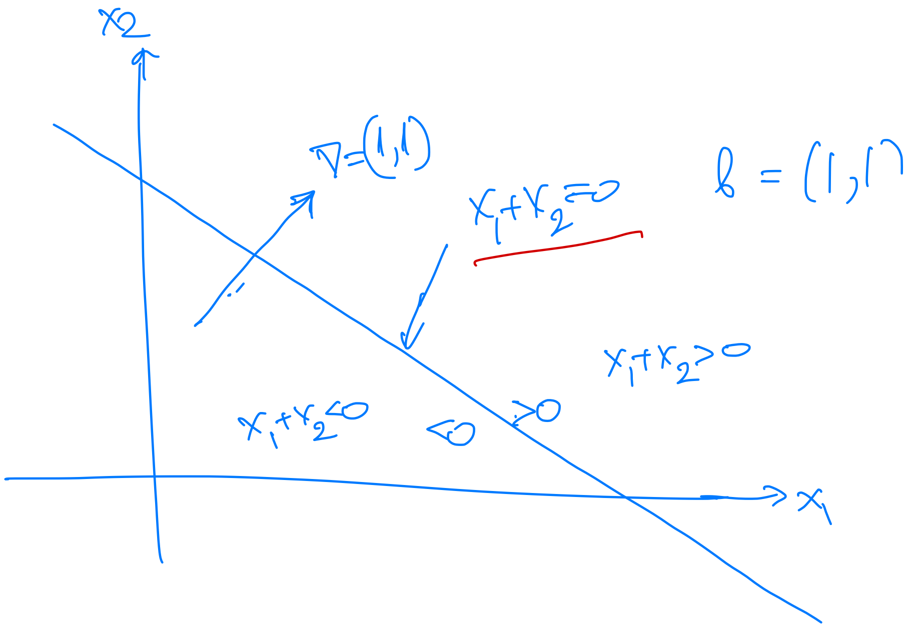
Separating hyp: $L=\{x: b_0 + x^T b = 0\}$

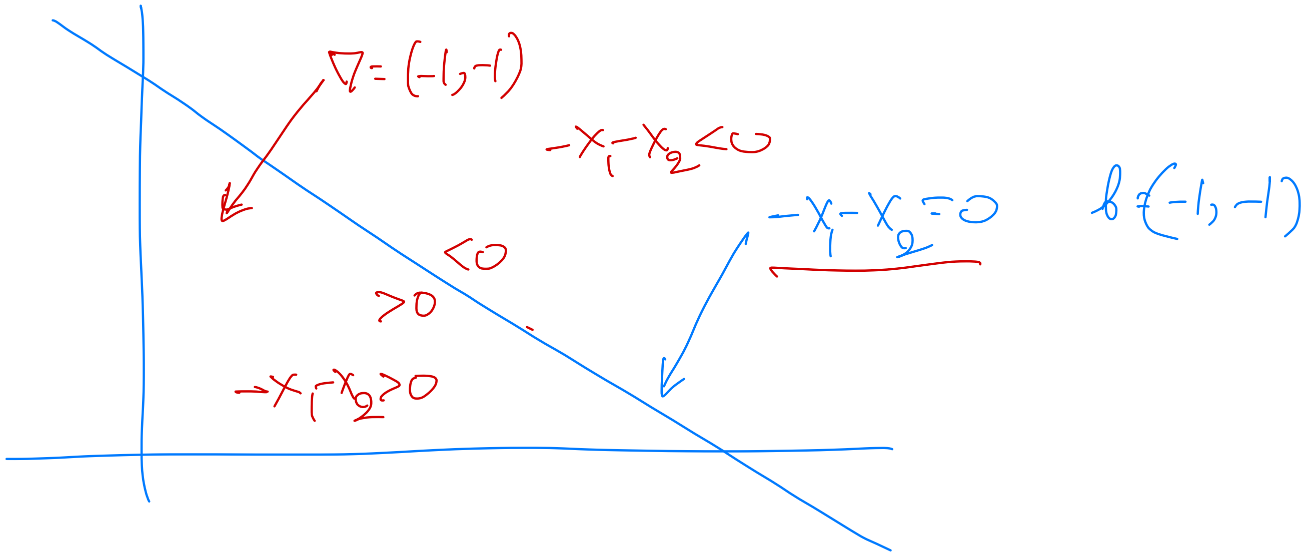
$\forall i$: $y_i = 1 \iff b_0 + x_i^T b > 0$ $\forall i$ — N περιορισμοί

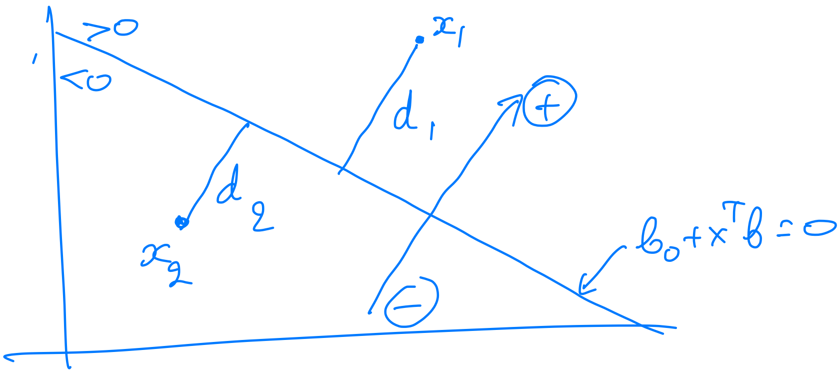
$y_i = -1 \iff b_0 + x_i^T b < 0$.



d = απόσταση x_0 από L.

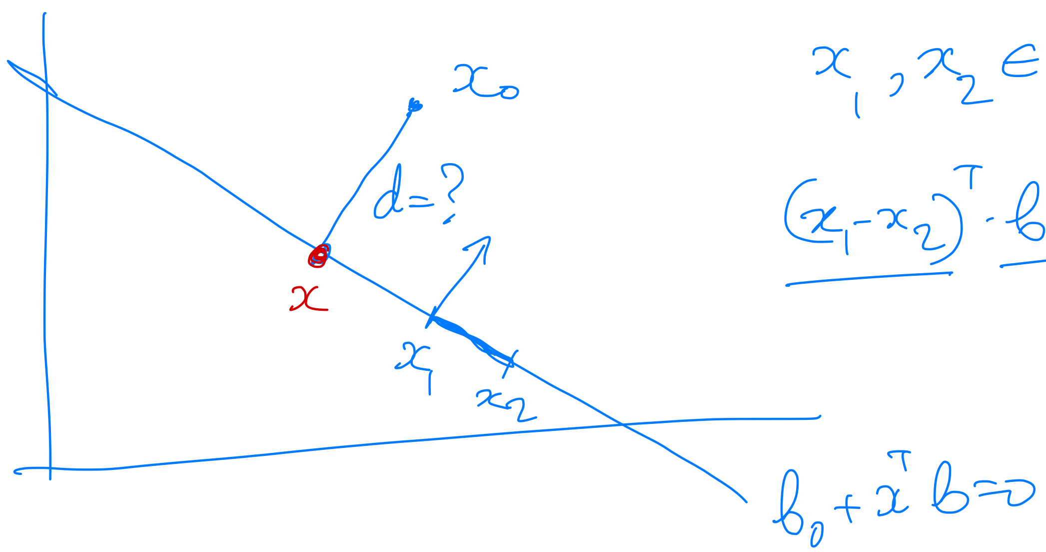






$$\text{signed distance} = \begin{cases} d & \text{αν } b_0+x^T b>0 \\ -d & \text{αν } b_0+x^T b<0 \end{cases}$$

d ευκλείδεια απόσταση



$x_1, x_2 \in L$

$(x_1-x_2)^T \cdot b = 0$

Πρόβλημα βελτιστοποίησης

$\forall x \in L: \quad b_0+x^T b=0$

$$\|x-x_0\|^2 = (\text{απόσταση } x_0 \to x)^2$$

$$\min_{x\in L} \left\{ \|x-x_0\|^2 \right\}$$

$$d^2(x_0) = \min_x \; (x - x_0)^T \cdot (x - x_0)$$
$$b_0 + x^T b \;\; = 0$$

Lagrangean : $L(x, \mu) = (x - x_0)^T \cdot (x - x_0) - \mu \left(b_0 + x^T b \right)$

$$\nabla_x L = 2(x - x_0) - \mu b = 0 \Rightarrow x^* - x_0 = \underline{\underline{\frac{\mu}{2} \cdot b}}$$

Για b_0 : $x \in L$: $b_0 + x^T b = 0 \Rightarrow$

$$\Rightarrow b_0 + \left(x_0 + \frac{\mu}{2} b \right)^T \cdot b = 0$$

$$\Rightarrow b_0 + x_0^T b = -\frac{\mu}{2} b^T b \Rightarrow$$

$$\Rightarrow \frac{\mu^*}{2} = -\frac{b_0 + x_0^T b}{\|b\|^2} \Rightarrow$$

$$\Rightarrow x^* - x_0 = -\frac{b_0 x_0^T b}{\|b\|^2} \cdot b = \alpha \cdot b$$

$$\Rightarrow d^2 = (x^* - x_0)^T \cdot (x^* - x_0) = \alpha^2 \cdot b^T b = \alpha^2 \|b\|^2$$

$$= \frac{\left(b_0 + x^T b \right)^2}{\|b\|^4} \|b\|^2 = \frac{\left(b_0 + x_0^T b \right)^2}{\|b\|^2}$$

$$d = \frac{|b_0 + x_0^T b|}{\|b\|}, \quad \text{signed distance} = \frac{b_0 + x_0^T b}{\|b\|}$$

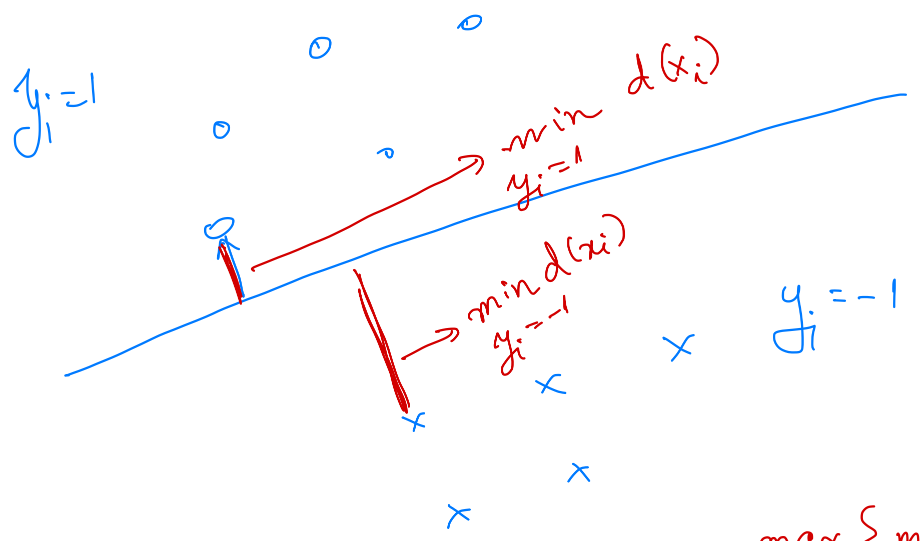
Αν θέσουμε $f(x) = b_0 + x^T b$

<u>classifier</u> $\quad G(x) = \text{sign}\, f(x) = \begin{cases} 1, & f(x) > 0 \\ -1, & f(x) < 0 \end{cases}$

$$\text{signed distance} = \frac{f(x_0)}{\|f'(x_0)\|} = \frac{f(x_0)}{\|\nabla_{x_0} f(x_0)\|}$$

Βέλτιστο διαχωρίζον υπερεπίπεδο

<u>Κριτήριο</u>: Μεγιστοποιεί την απόσταση από το κοντινότερο σημείο για κάθε κλάση



$$\max_{b_0, b_1} \left\{ \min_{i \in N} d(x_i) \right\}$$

Το πρόβλημα βελτιστοποίησης γράφεται:

$$\max_{b_0,\, b} \; M$$

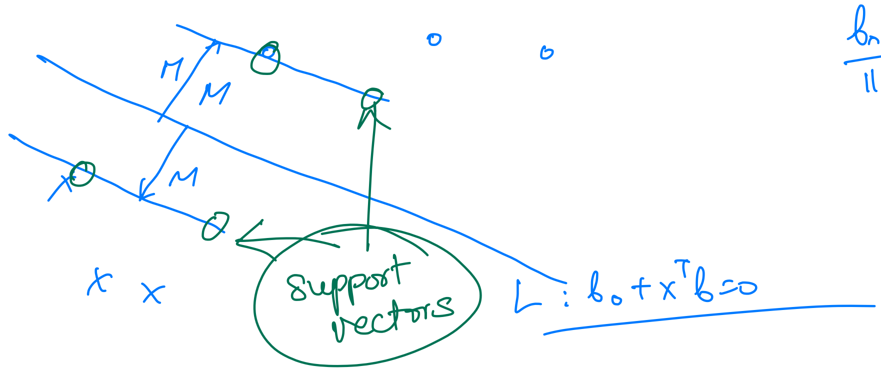
$$\text{s.t.} \quad \|b\|^2 = 1$$

γραμμ. περ. $\quad y_i(b_0 + x_i^T b) \geq M \quad \forall i = 1, \ldots, N \iff$ αν $y_i = 1 \Rightarrow b_0 + x_i^T b \geq M$, αν $y_i = -1 \Rightarrow b_0 + x_i^T b \leq -M$

$\rightarrow$ SV Classifier

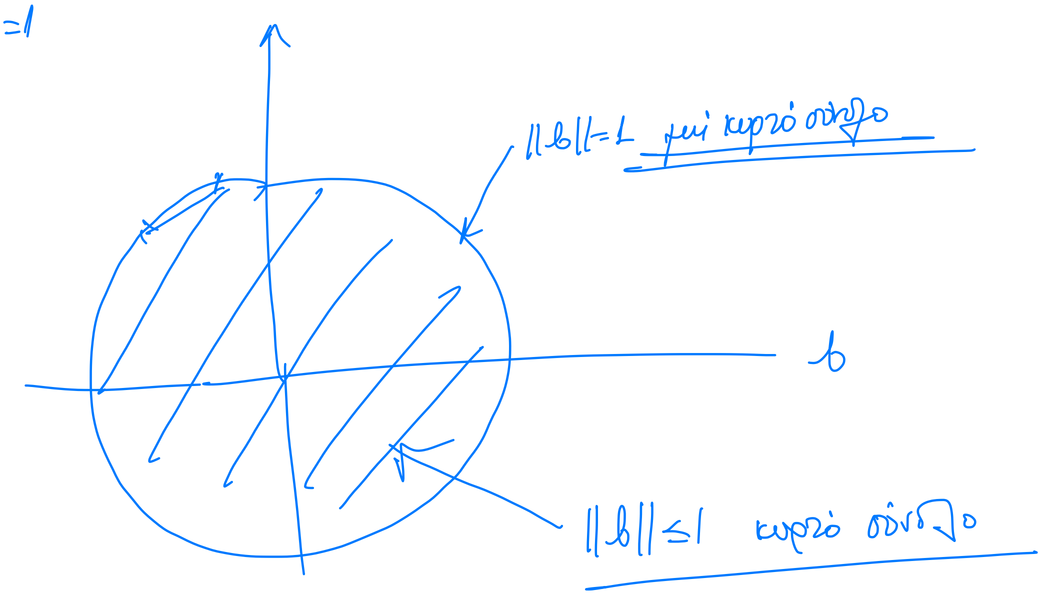
απόσταση x_i από L $\geq M$

$$\frac{b_0 + x_i^T b \geq M}{\|b\|}$$



$y_i(b_0 + x_i^T b)$ = ευκλ. απόσταση του x_i από L $(\|b\| = 1)$

$\|b\| = 1$



Μετ/σμός σε πρόβλημα κυρτού προγ/σμού

Αν αγνοήσουμε τον περιορισμό $\|b\|=1$

$$\max_{b_0, b} M$$

$$\text{s.t.} \quad y_i \frac{b_0 + x_i^T b}{\|b\|} \geq M \Rightarrow y_i(b_0 + x_i^T b) \geq M\|b\|$$

Όμως το $\|b\|$ είναι αυθαίρετο

Επομένως παίρνουμε $\|b\| = \frac{1}{M} \Rightarrow M = \frac{1}{\|b\|}$

όχι κυρτή

$$\max \frac{1}{\|b\|} \qquad \Longleftrightarrow \qquad \boxed{\begin{aligned} &\min \|b\| \\ &y_i(b_0 + x_i^T b) \geq 1 \end{aligned}}$$

$$y_i(b_0 + x_i^T b) \geq 1$$

Τελική μορφή

$$\boxed{\begin{aligned} &\min \frac{1}{2}\|b\|^2 \\ &\text{s.t.} \quad y_i(b_0 + x_i^T b) \geq 1 \quad \forall i = 1, \ldots, N \end{aligned}}$$