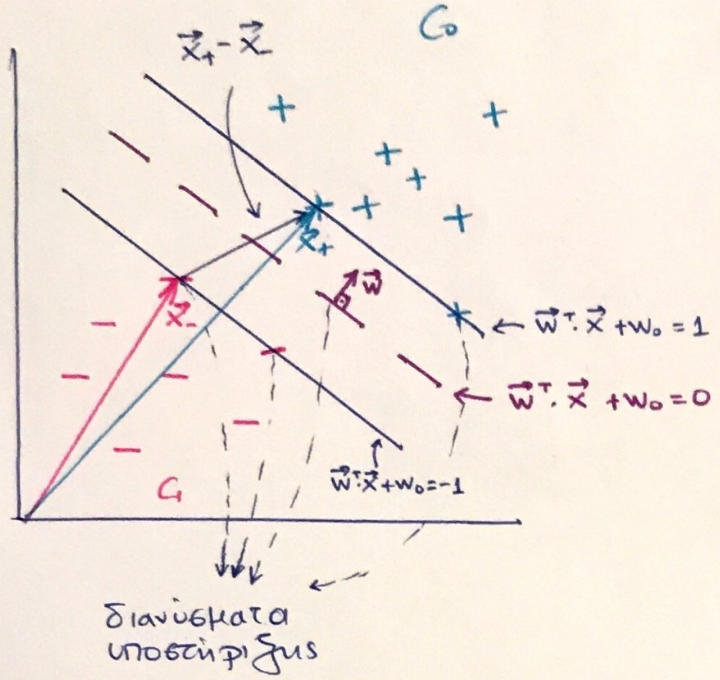

$\vec{w}^T \cdot \vec{x}_i + w_0 \geq 1$ για $x_i \in C_0$ (1)

$\vec{w}^T \cdot \vec{x}_i + w_0 \leq -1$ για $x_i \in C_1$ (2)

$$d_i = \begin{cases} +1, & x_i \in C_0 \\ -1, & x_i \in C_1 \end{cases}$$

Πολλαπλασιάζοντας τις (1), (2) με d_i:

$d_i (\vec{w}^T \cdot \vec{x}_i + w_0) \geq 1 \Rightarrow$

$d_i (\vec{w}^T \cdot \vec{x}_i + w_0) - 1 \geq 0$ (3)

Για τα $\vec{x}_i$ στο πάνω και κάτω όριο του περιθωρίου (έστω $\vec{x}_+$ και $\vec{x}_-$ αντίστοιχα):

$d_i (\vec{w}^T \cdot x_i + w_0) - 1 = 0$ (4)

Το πλάτος γ του περιθωρίου ισούται με:

$$\frac{\vec{w}^T}{\|\vec{w}\|} \cdot (\vec{x}_+ - \vec{x}_-) = \frac{1}{\|\vec{w}\|} \cdot (\vec{w}^T \cdot \vec{x}_+ - \vec{w}^T \cdot \vec{x}_-) \overset{(4)}{=\!=} \frac{1}{\|\vec{w}\|} \cdot [(1 - w_0) - (-1 - w_0)] = \frac{2}{\|\vec{w}\|} \quad (5)$$

Επιδιώκουμε να μεγιστοποιήσουμε το γ:

$$\max \gamma \rightarrow \max \frac{2}{\|\vec{w}\|} \longrightarrow \min \frac{\|\vec{w}\|^2}{2}.$$

Μέθοδος Lagrange

$$\mathcal{L} = \frac{1}{2} \|\vec{w}\|^2 - \sum_i \lambda_i [d_i (\vec{w}^T \cdot \vec{x}_i + w_0) - 1] \quad \text{με } \lambda_i \geq 0 \qquad (6)$$

Συνθήκες Karush-Kuhn-Tucker

$$\frac{\partial \mathcal{L}}{\partial w_0} = 0 \Rightarrow -\sum_i \lambda_i d_i = 0 \Rightarrow \boxed{\sum_i \lambda_i d_i = 0} \quad (7)$$

$$\frac{\partial \mathcal{L}}{\partial \vec{w}} = 0 \Rightarrow \vec{w} - \sum_i \lambda_i d_i \vec{x}_i = 0 \Rightarrow \boxed{\vec{w} = \sum_i \lambda_i d_i \vec{x}_i} \quad (8)$$

$$(6) \xrightarrow{(7), (8)} \mathcal{L} = \frac{1}{2} \left(\sum_i \lambda_i d_i \vec{x}_i^T\right) \cdot \left(\sum_j \lambda_j d_j \vec{x}_j\right) - \left(\sum_i \lambda_i d_i \cdot \sum_j \lambda_j d_j \vec{x}_j^T \cdot \vec{x}_i\right) -$$

$$- w_0 \cdot \cancelto{0}{\sum_i \lambda_i d_i} + \sum_i \lambda_i \Rightarrow$$

$$\mathcal{L} = \sum_i \lambda_i - \frac{1}{2} \sum_i \sum_j \lambda_i \lambda_j d_i d_j \vec{x}_i^T \cdot \vec{x}_j .$$