Théorème : Soit le train set

$$S = \{ (\vec{x}_k, y_k) \in \mathbb{R}^M \times \{-1, 1\} : k = 1, \ldots, K \}.$$

Supposons que :

(i) $\exists R > 0$ t.q. $\| \vec{x}_k \| \leq R$, $\forall k = 1, \ldots, K$

(ii) $\exists \vec{\hat{\omega}} \in \mathbb{R}^{M+1}$ et $\gamma > 0$ t.q.

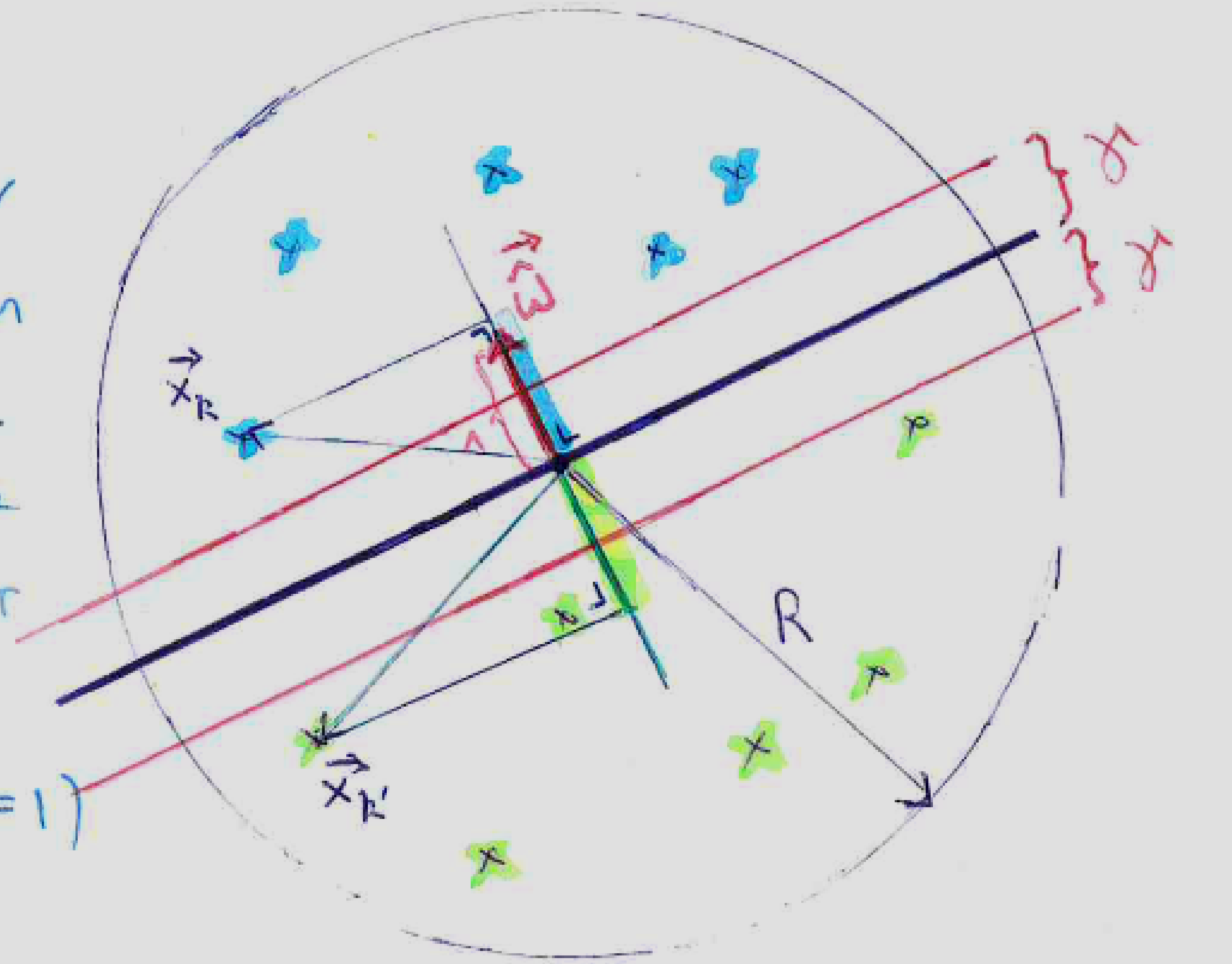
$\| \vec{\hat{\omega}} \| = 1$ et $y_k \cdot (\vec{\hat{\omega}}^T \cdot \vec{x}_k) \geq \gamma$, $\forall k = 1, \ldots, K$.

Alors l'algorithme d'entraînement du perceptron converge en au plus R^2 / γ^2 itérations.

Rappel :

$\vec{\hat{\omega}}^T \cdot \vec{x}_k$ est la longueur de la proj. orthogonale de $\vec{x}_k$ sur $\vec{\hat{\omega}}$

(car $\| \vec{\hat{\omega}} \| = 1$)



segment de longueur $y_k \cdot (\vec{\hat{\omega}}^T \cdot \vec{x}_k) \geq \gamma$

segment de longueur $y_{k'} (\vec{\hat{\omega}}^T \cdot \vec{x}_{k'}) \geq \gamma$

Preuve : On omet la notation vectorielle.

Soit w^k le vecteur des paramètres à l'itération k, avec $w^0 = 0$.

(1) On montre par recurrence que

$w^{k^T} \cdot \hat{w} \geq k \cdot \gamma$, pour toute itération k. $(\bullet)$

(i) Soit $k = 0$. On a :

$$w^{0^T} \hat{w} = 0^T \cdot \hat{w} = 0 = 0 \cdot \gamma = k \cdot \gamma$$

(ii) Supposons $(\bullet)$ vraie pour tout $k' \leq k$, et supposons que l'élément i soit mal classifié à l'itération $k+1$:

$$w^{k+1^T} \cdot \hat{w} \underset{\text{algo}}{=} (w^k + \overset{\in \mathbb{R}}{y_i} \cdot x_i)^T \cdot \hat{w}$$

$$= w^{k^T} \hat{w} + y_i \cdot x_i^T \hat{w}$$

$$\underset{\substack{\text{hyp d'ind.} + \\ \text{hyp du thm}}}{\geq} k \cdot \gamma + \gamma = (k+1) \cdot \gamma \qquad \square$$

On a alors :

$$\|w^k\| = \|w^k\| \cdot 1 \underset{\substack{\text{hyp} \\ \text{thm}}}{=} \|w^k\| \cdot \|\hat{w}\|$$

$$\underset{\substack{(a^T b = \|a\| \cdot \|b\| \cos\theta \\ \leq \|a\| \cdot \|b\|)}}{\geq} w^{k^T} \cdot \hat{w} \underset{(\bullet)}{\geq} k \cdot \gamma \qquad (*)$$

(2) On montre par récurrence que

$\|w^k\|^2 \leq k \cdot R^2$ pour toute itération k. $(**)$

(i) Pour $k=0$, on a :

$\|w^0\|^2 = \|0\|^2 = 0 \cdot R^2$

(ii) Supposons $(**)$ vraie pour tout $k' \leq k$, et supposons que l'élément x_i est mal classifié à l'itération $k+1$:

$$\|w^{k+1}\|^2 \underset{\text{algo}}{=} \|w^k + y_i \cdot x_i\|^2$$

$$= \|w^k\|^2 + \underbrace{2\, y_i \cdot w^{k^T} \cdot x_i}_{<0} + \underbrace{\overset{\overset{1}{\shortparallel}}{y_i^2}\, \|x_i\|^2}_{\leq R^2}$$

car x_i mal classifié et hyp. du thm
$$\leq \|w^k\|^2 + R^2 \underset{\text{hyp ind}}{\leq} k \cdot R^2 + R^2 = (k+1) \cdot R^2 \quad \square$$

Pour toute itération k (non terminale) de l'algo, on a donc par $(*)$ et $(**)$

$$k^{\cancel{2}} \cdot \gamma^2 \underset{(*)}{\leq} \|w^k\|^2 \underset{(**)}{\leq} \cancel{k} \cdot R^2$$

Et donc : $k \leq \dfrac{R^2}{\gamma^2}$

Ainsi, si k est une itération non terminale, alors $k \leq \dfrac{R^2}{\gamma^2}$

$\blacksquare$