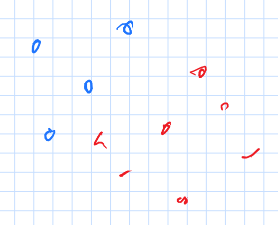
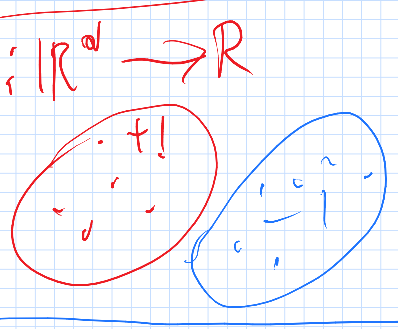
12/4/18

Perceptron

$\mathbb{R}^d$ d large

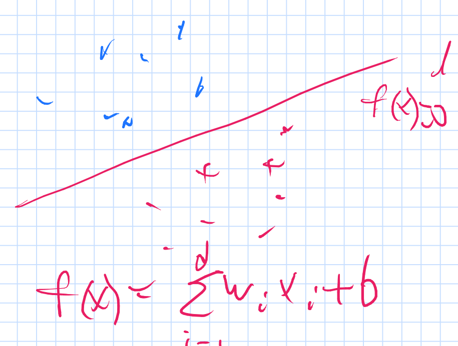


$f: \mathbb{R}^d \to \mathbb{R}$



Occam's razor

Find the "simplest" function that classifies the samples correctly.



$f(x) = \sum_{i=1}^{d} w_i x_i + b$

$\text{sign}(f(x))$

$p_1, \dots, p_t$ n positive points

$n_1, \dots, n_t$ negative points

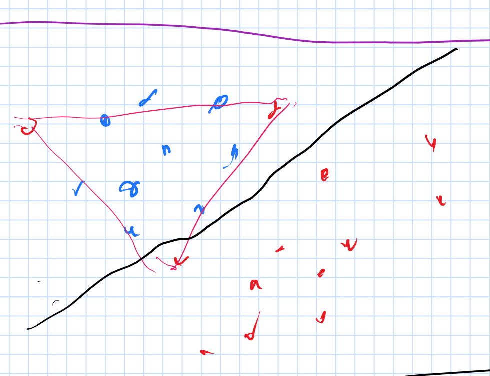
$(w_1, w_2, \dots, w_d)$

$\forall i \quad \langle w, p_i \rangle + b \geq \gamma$

$\forall j \quad \langle w, n_j \rangle + b \leq -\gamma$

Linear program

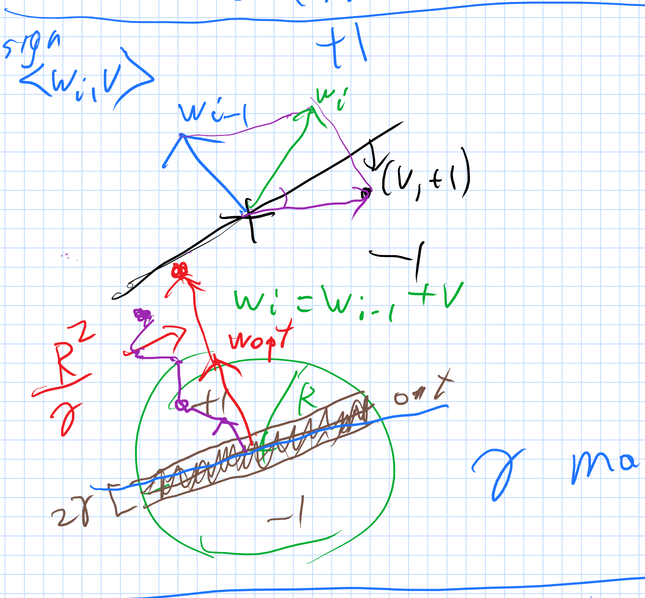
γ tiny constant.



$(v_1, y_1), \dots, (v_n, y_n)$ $\quad v_i \in \mathbb{R}^n$ $\quad y_i \in \{-1, +1\}$

Looking for a seprating hyperplane through the origin.

$w_0 \leftarrow 0$
$i = 0$
while there is a misclassified example (v, y) do:
// $\text{sign}(\langle w_{i-1}, v \rangle) \neq y$
$w_i \leftarrow w_{i-1} + yv$
return $f(x) = \langle w, x \rangle$ as the classifier



Theorem The perceptron alg performs $\left(\frac{R}{\gamma}\right)^2$ misclassifications mistakes.

No dependency on d!
No dependency on n.

Proof potential function

$\alpha_i = \left\| w_i - \frac{R^2}{\gamma} w_{opt} \right\|^2$

$\|w_{opt}\| = 1$

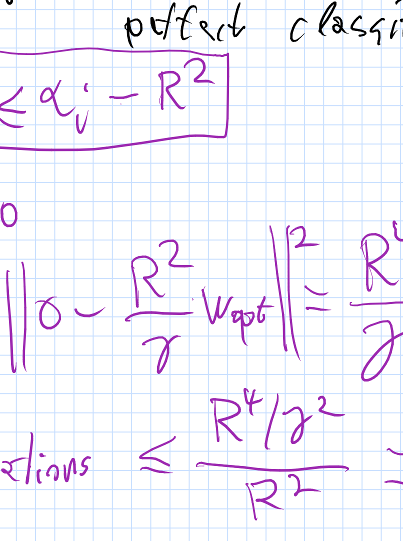
w_{opt} is the normal of the perfect classifier.

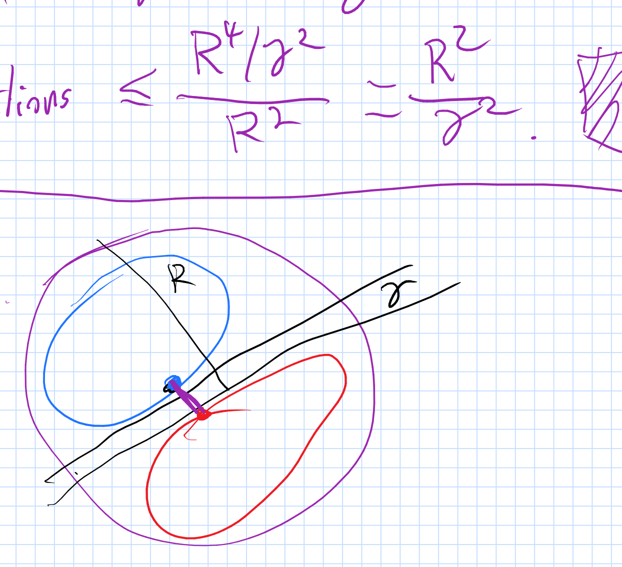
$\boxed{\alpha_{i+1} \leq \alpha_i - R^2}$

$w_0 = 0$

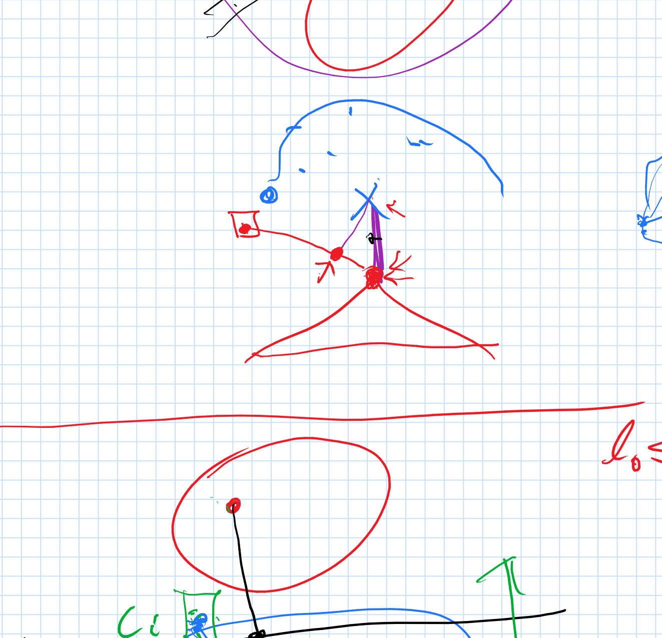
$\alpha_0 = \left\| 0 - \frac{R^2}{\gamma} w_{opt} \right\|^2 = \frac{R^4}{\gamma^2}$

$\#\text{ iterations} \leq \frac{R^4/\gamma^2}{R^2} = \frac{R^2}{\gamma^2}$. ▨





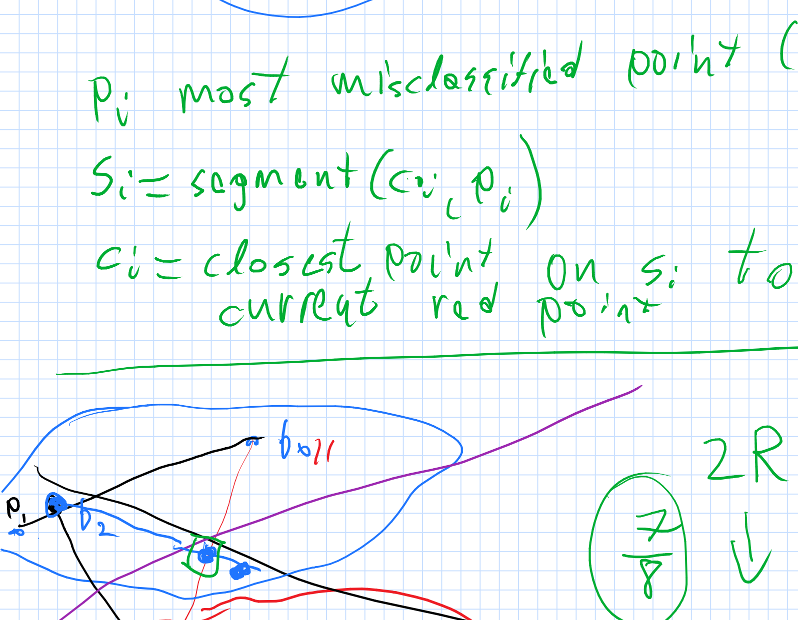
$l_0 \leq 2R$

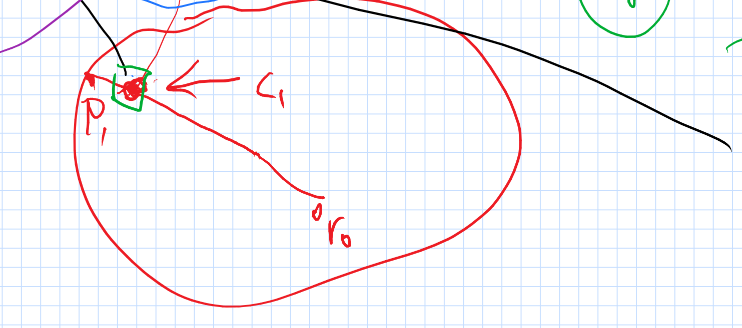


p_i most misclassified point (blue)

s_i = segment (c_{i-1}, p_i)

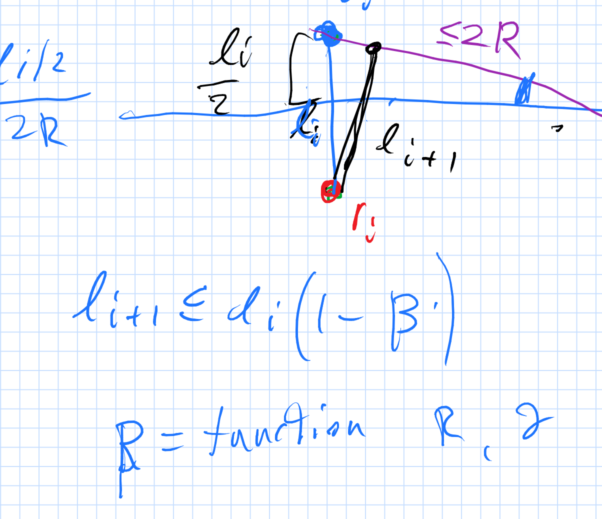
c_i = closest point on s_i to the current red point





$l_{i+1} \leq l_i (1 - \beta)$

β = function R, γ

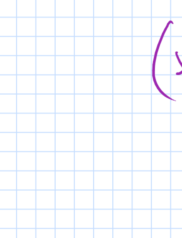


$(x_1, x_2, \dots, x_d) \xrightarrow{\text{lifting}}$

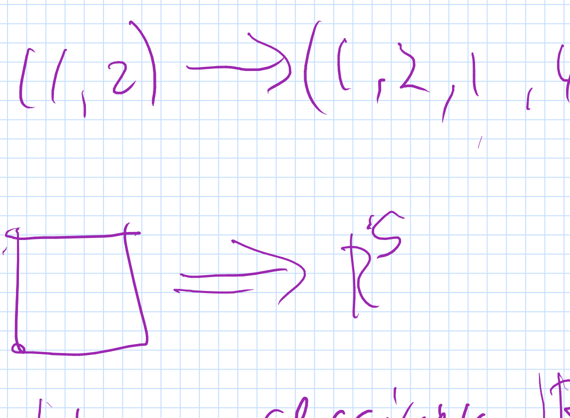
$(x_1, x_2, \dots, x_d, x_1^2, x_2^2, \dots, x_d^2,$
$x_1 x_2, \dots$
$x_7^3 x_5^2 \dots$

$(x, y) \to (x, y, x^2, y^2, xy)$

$(1, 2) \to (1, 2, 1, 4, 2)$

 $\Rightarrow \mathbb{R}^5$

Linear classifier $\mathbb{R}^5 \Rightarrow$ quadratic polynomial in the plane



$\langle w_i, v \rangle$

(original point) $\xrightarrow{\text{lifting}}$ (______)

(______)
w

$\langle x, y \rangle$

|| ||