LEARNING $\underline{w}$ WITH PERCEPTRON (CH. 4.1.7)

INPUT SPACE → FEATURE SPACE $\quad \underline{x} \to \phi(\underline{x})$

$M = 2$

$$y(\underline{x}) = f(\underline{w}^T \phi(\underline{x}))$$

f: $+1 \geq 0$ C_1; $-1 < 0$ C_2

$\underline{w}^T \phi(\underline{x}) > 0$ C_1

< 0 C_2

t_n: $+1$ C_1; -1 C_2

CORRECT CLASSIFICATION:

$$\underline{w}^T \phi(\underline{x}_n)\, t_n > 0 \quad (!)$$

LEARNING:

$$\text{MIN}\left(-\underline{w}^T \phi(\underline{x}_n)\, t_n\right)$$

$\uparrow$ $\sum_{n \in M}$ MISSCLASSIFIED

$$E_p(\underline{w}) = -\sum_M \underline{w}^T \phi(\underline{x}_n)\, t_n$$

GRADIENT DESCENT:
$$\underline{w}^{(\tau+1)} = \underline{w}^{(\tau)} - \eta \nabla_{\underline{w}} E_p(\underline{w})$$
(η: LEARNING RATE)

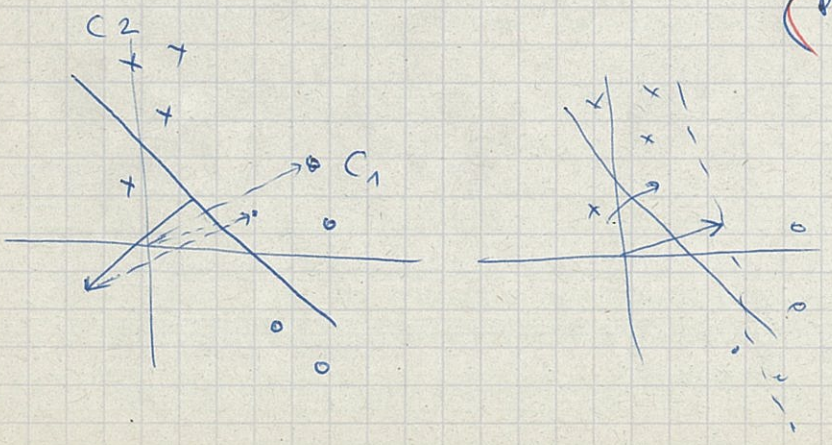
SEQUENTIAL FOR $\underline{x}_n$:
$$\underline{w}^{(\tau+1)} = \underline{w}^{(\tau)} + \eta\, \phi(\underline{x}_n) \cdot t_n$$

- $\underline{x}_n$ OK, NO
- — MISSCLASSIFIED

FOR
- C_1: $\underline{w} + \phi(\underline{x}_n)$
- C_2: $\underline{w} - \phi(\underline{x}_n)$

PERCEPTRON CONVERGENCE THEOREM: IF SOLUTION EXISTS (CLASSES LINEARLY SEPARABLE) LEARNING CONVERGENT TO THE SOLUTION IN FINITE NUMBER OF STEPS

(PROBLEM: SLOW OR NON SEPARABLE ?)



(FIG. 4.7 BISHOP)