classification with linear Model with a hard threshold

① linear model for classification

x_1	x_2	Y
2	2	1
3	3	1
4	1	0
5	0	0

0 : underground explosion

1 : earth-quake.
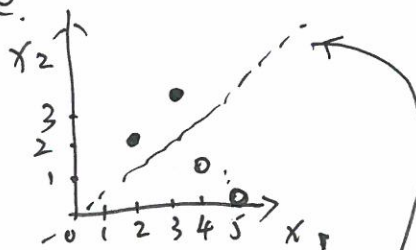
$y = ax + b$.

$x_2 = W_1 x_1 + W_0 \rightarrow \boxed{x_2 = x_1 - 0.5 x_0}$

$(W_2 = 1, \quad W_1 = 1, \quad W_0 = -0.5)$

x_0 dummy inp.
|
1

$(\frac{x_2}{x_1})$.

$\Rightarrow \boxed{x_2 - x_1 + 0.5 x_0 = 0.} \longrightarrow$ linear model

$(2, 2) \quad \underline{2 - 2} + \underline{0.5 \times 1} = 0.5 > 0 \quad \Rightarrow 1$

$(4, 1) \quad \xcancel{4 \ 1} \quad 1 - 4 + 0.5 \times 1 = -2.5 < 0 \Rightarrow 0$

② How to learn ③ parameters?   why 3 : $\rightarrow \begin{array}{|ccc|} x_2 & x_1 & x_0 \\ W_2 & W_1 & W_0 \end{array}$ T(w·x).

$d + 1$ parameters
↑
number of features

$\hat{y} = T(\underline{W_0 x_0} + \underline{W_1 x_1} + \underline{W_2 x_2})$ $w \cdot x$

☹ not differentiable

☺ perceptron learning Rule
(== stochastic GD)

$W_i \leftarrow W_i + \alpha \cdot (\underline{y - \hat{y}}) \times x_i$

case 1: $\hat{y} == y$. $W_i \leftarrow W_i$
case 2: $\hat{y} = 0, y = 1$. $W_i \uparrow \leftarrow W_i + pos.$
case 3: $\hat{y} = 1, y = 0$ $W_i \downarrow \leftarrow W_i + neg.$

③ $W_0 = 0.5$ $W_1 = 0.5$ $W_2 = 0.5$ $\alpha = 0.1$

iter	x_0 x_1 x_2 y	
1	(1.2.2.1)	$1 \times 0.5 + 2 \times 0.5 + 2 \times 0.5 > 0 \to 1 = y$
2	(1.3.3.1)	$1 \times 0.5 + 3 \times 0.5 + 3 \times 0.5 > 0 \to 1 = y$
3	(1,4.1,0)	$0.5 \times 1 + 0.5 \times 4 + 0.5 \times 1 > 0 \to 1 \ne y$

$$W_0 = 0.5 + 0.1 \times (0-1) \times 1 = 0.4$$
$$W_1 = 0.5 + 0.1 \times (0-1) \times 4 = 0.1$$
$$W_2 = 0.5 + 0.1 \times (0-1) \times 1 = 0.4$$

4	(1.5. 0. 0)	$0.4 \times 1 + 0.1 \times 5 + 0.4 \times 0 = 0.9 > 0 \to 1 <> y$

$$W_0 = 0.4 + 0.1 \times (0-1) \times 1 = 0.3$$
$$W_1 = 0.1 + 0.1 \times (0-1) \times 5 = -0.4$$
$$W_2 = 0.4 + 0.1 \times (0-1) \times 0 = 0.4$$

5	(1.2.2.1)	$0.3 \times 1 + -0.4 \times 2 + 0.4 \times 2 \geq 0 \to 1 = y$
6	(1.3.3.1)	$- \quad - \quad - \quad - \quad - \quad > 0 \to 1 = y$
7	(1.4.1.0)	$0.3 \times 1 + -0.4 \times 4 + 0.4 \times 1 \leq 0 \to 0 = y$
8	(1.5, 0, 0)	$0.3 \times 1 + -0.4 \times 5 + 0.4 \times 0 < 0 \to 0 = y$

} zero error convergence

$$0.3 + -0.4 x_1 + 0.4 x_2 = 0$$
$$3 + -4 x_1 + 4 x_2 = 0$$

$$\boxed{x_2 = x_1 - 0.75}$$

④convergence

"Artificial Intelligence: A Modern Approach" by

Stuart. Russell, Peter Norvig.

check: section 18.6.3 (p725. third Edition)
figure 18.16