$\mathbb{X}, \mathbb{Y}$ $\quad f: \mathbb{X} \to \mathbb{Y}$

$X \subset \mathbb{X}$ $\quad \{(x_i, y_i) \in X \times Y\}$

$\hat{f}$

1) $|Y| = const$ classification

2) $|Y| = |\mathbb{R}|$ $[0, 130]$

3) $|Y| = const$ clustering

$\{ x_i \in X \}$

4) $x_i \sim N(0,1)$, $x \in \mathbb{R}$ $\quad y_i \in \mathbb{R}^{h \times w \times 3}$

$\hat{t} = \hat{f}(X, \theta)$ — model

$\underset{\theta}{\arg\min}\ L(\hat{t}, t_i)$

$\hat{f}(x, \theta) = \theta_0 x + \theta_1$

1) train / val / test split.

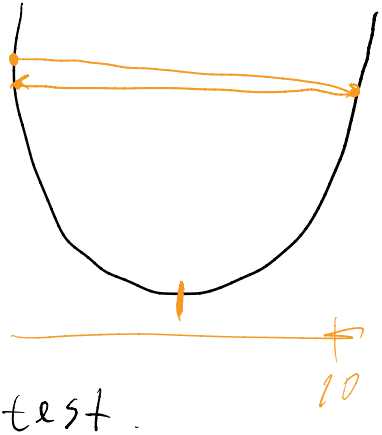
2) loss function: $MSE = \|\hat{t}_\theta - t_i\|_2$

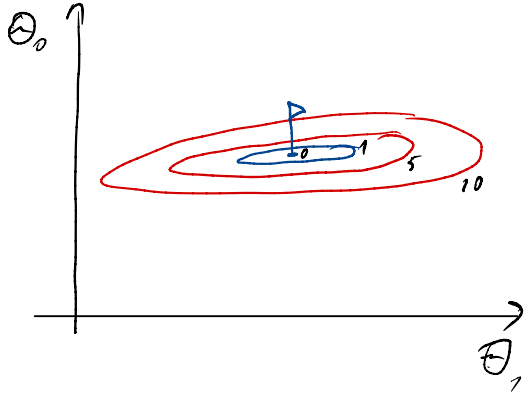
$MAE = \|\hat{t}_\theta - t\|_1$

3). $L \to \min$

I) $\theta_0 \sim N(0, 1)$

II) $\theta_i = \theta_{i-1} - \alpha \nabla_\theta L(\hat{f}_\theta(X), f(X))$





III) test.

$\hat{f}(x, \theta) = \sigma(\theta^T x)$