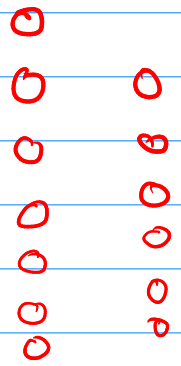
$(x_1, t_1)\ (x_2, t_2) \cdots (x_N, t_N)$

$y = w_0 + w_1 x^1 + w_2 x^2 \cdots + w_q x^q$

$$E(w) = \sum_{n=1}^{N} (t_n - y(x_n, w))^2$$

Squared sum of all weights $\leftarrow \|w\|^2 \leq \eta \Rightarrow$

$$J(\theta) = \sum_{n=1}^{N} (t_n - y(x_n, w))^2 + \frac{\alpha}{2} \|w\|^2$$



$$\frac{1}{2} (\hat{y}_n - y)^2 + \frac{\alpha}{2} \|w\|^2$$

↓ not biases

bias → extra node with '1' weight

We regularize weights, but not biases normally

(it is to provide a bias; it is hardwired!

can be updated but rarely is