$$J(w) = \frac{1}{2N} \sum_{n=1}^{N} \left(w^T x^{(n)} - t_n \right)^2 + \frac{\lambda}{2} \|w\|^2$$

$w = [w_0, w_1, \dots, w_M]^T$

$x^{(n)} = [1, x_n, x_n^2, \dots, x_n^M]$

$w^T x^{(n)} = w_0 + w_1 x_n + \dots + w_M x_n^M$

$\|w\|^2 = \sum_{j=0}^{M} w_j^2$

λ → hyper-parameter. ≥ 0

$$\hat{w} = \arg\min_{w} J(w) \qquad \frac{\partial J}{\partial w} = 0$$

$\ln \lambda = -18 \Rightarrow \lambda = e^{-18} = \frac{1}{e^{18}}$

e → 2.71… Euler's number.

$e = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$ $\qquad e^{\pi i} = -1.$

Why not $\hat{w}, \hat{\lambda} = \arg\min_{w, \lambda} J(w, \lambda) \Rightarrow \hat{\lambda} = 0$!

Dataset D