If λ is very large,

$$J(\theta) = \frac{1}{2m}\left[\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right) + \lambda \sum_{j=1}^{n} \theta_j^2\right]$$

$$h_\theta(x) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_4 x^4$$

for very large $\lambda \Rightarrow \theta_1 \approx 0, \theta_2 \approx 0 \cdots, \theta_n \approx 0$

$\therefore h_\theta(x) = \theta_0 \Rightarrow$ constant line $\parallel$r to x axis

Underfit or 'high bias'

Regularization linear regression

$$J(\theta) = \frac{1}{2m}\left[\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)^2 + \lambda \sum_{j=1}^{n} \theta_j^2\right]$$

$$\min_\theta J(\theta)$$

Grad. Desc.

Repeat {

$$\theta_0 := \theta_0 - \frac{\alpha}{m} \sum_{i=1}^{m}\left(\left(h_\theta(x^{(i)}) - y^{(i)}\right) x_0^{(i)}\right)$$

$$\theta_j := \theta_j - \frac{\alpha}{m} \sum_{i=1}^{m}\left(\left(h_\theta(x^{(i)}) - y^{(i)}\right) x_j^{(i)}\right) + \frac{\lambda}{m}\theta_j$$

}