

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

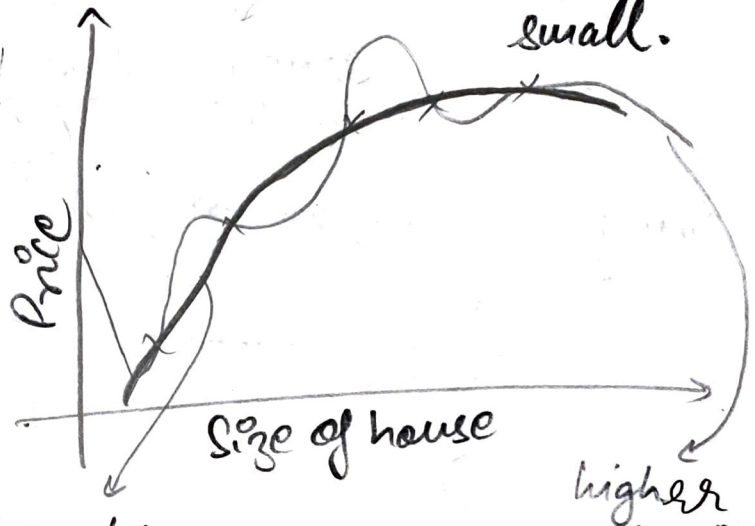
Regularization  
parameter

this term helps  
fitting training  
data well.

Regularization  
term

this term  
helps in keeping  
the parameters  
small.

$$\min_{\theta} J(\theta)$$



higher order  
hypo. func.  
with regularization  
(smoother and better)

higher  
order hypo. func.  
without regularization