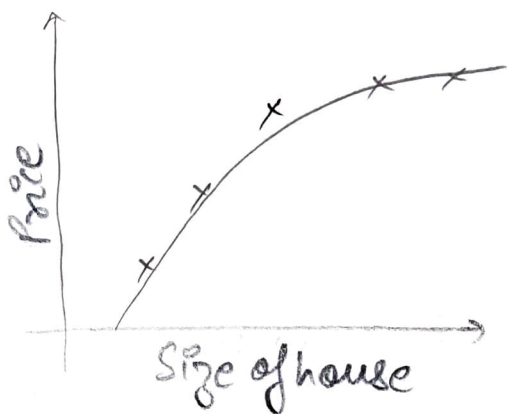


## Intuition



$$\theta_0 + \theta_1 x + \theta_2 x^2$$



$$\theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3$$

Suppose we penalize and make  $\theta_3$  &  $\theta_4$  really small.

$$\min_{\theta} \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2 + 1000\theta_3^2 + 1000\theta_4^2$$

$$\hookrightarrow \theta_3 \approx 0 \text{ \& \> } \theta_4 \approx 0 \Rightarrow \theta_0 + \theta_1 x + \theta_2 x^2 + 0 + 0$$

## Regularization

Small values for parameters  $\theta_1, \theta_2, \dots$

Or

→ Simpler hypothesis

→ Less prone to overfitting

(eg) Housing?

→ features  $\Rightarrow x_1, x_2, \dots, x_{100}$

→ Parameters  $\Rightarrow \theta_0, \theta_1, \theta_2, \dots, \theta_{100}$

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