

$$h_0(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 x^3$$

$$= \theta_0 + \theta_1 (\text{size}) + \theta_2 (\text{size})^2 + \theta_3 (\text{size})^3$$

$$x_1 = (\text{size}) ; x_2 = (\text{size})^2 ; x_3 = (\text{size})^3$$

③

3 different features

using single given feature
(size $\rightarrow x$)

POLYNOMIAL REGRESSION

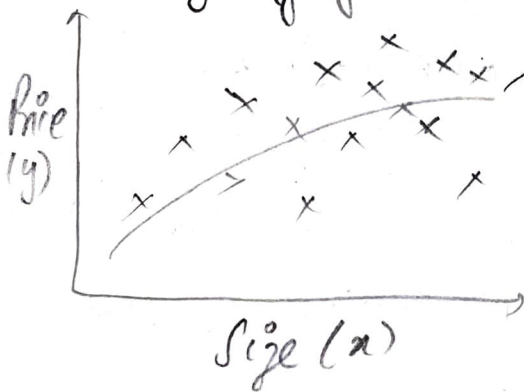
If, size: $1 \rightarrow 10^3$ size³ = $1 \rightarrow 10^9$

then, size² = $1 \rightarrow 10^6$

quite large range

use feature scaling
to improve gradient descent.

Choosing diff. features



$$h_0(x) = \theta_0 + \theta_1 x + \theta_2 \sqrt{x}$$

$$= \theta_0 + \theta_1 (\text{size}) + \theta_2 \sqrt{\text{size}}$$

using diff set of
features for same set
of training data