Feature engineering

$$f_{\vec{w},b}(\vec{x}) = w_1 x_1 + w_2 x_2 + b$$

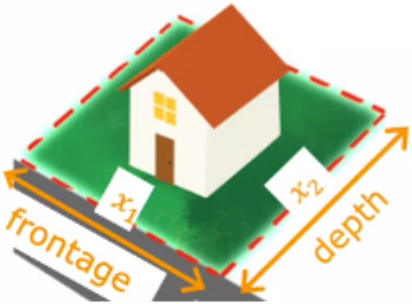
x_1: frontage, x_2: depth

$$area = frontage \times depth$$

$$x_3 = x_1 x_2$$

new feature

$$f_{\vec{w},b}(\vec{x}) = w_1 x_1 + w_2 x_2 + w_3 x_3 + b$$

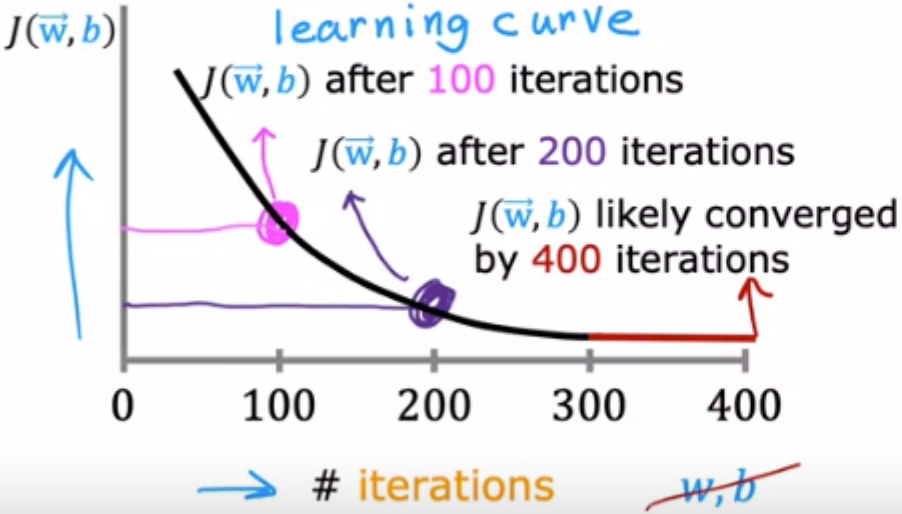


Feature engineering:
Using intuition to design new features, by transforming or combining original features.

Make sure gradient descent is working correctly

objective: $\min_{\vec{w},b} J(\vec{w},b)$

$J(\vec{w},b)$ should decrease after every iteration



iterations needed varies 30 1,000 100,000

Automatic convergence test

Let ε "epsilon" be 10^{-3}. 0.001

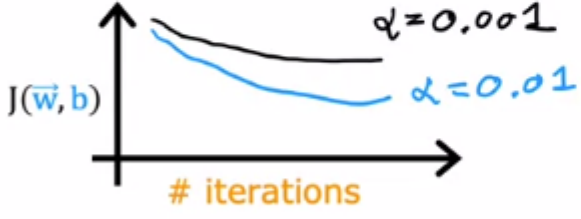
If $J(\vec{w},b)$ decreases by $\leq \varepsilon$ in one iteration, declare convergence.

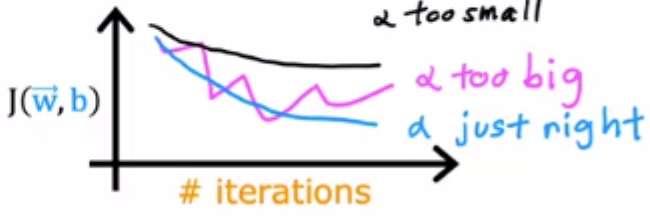
(found parameters $\vec{w}, b$ to get close to global minimum)

Values of α to try:

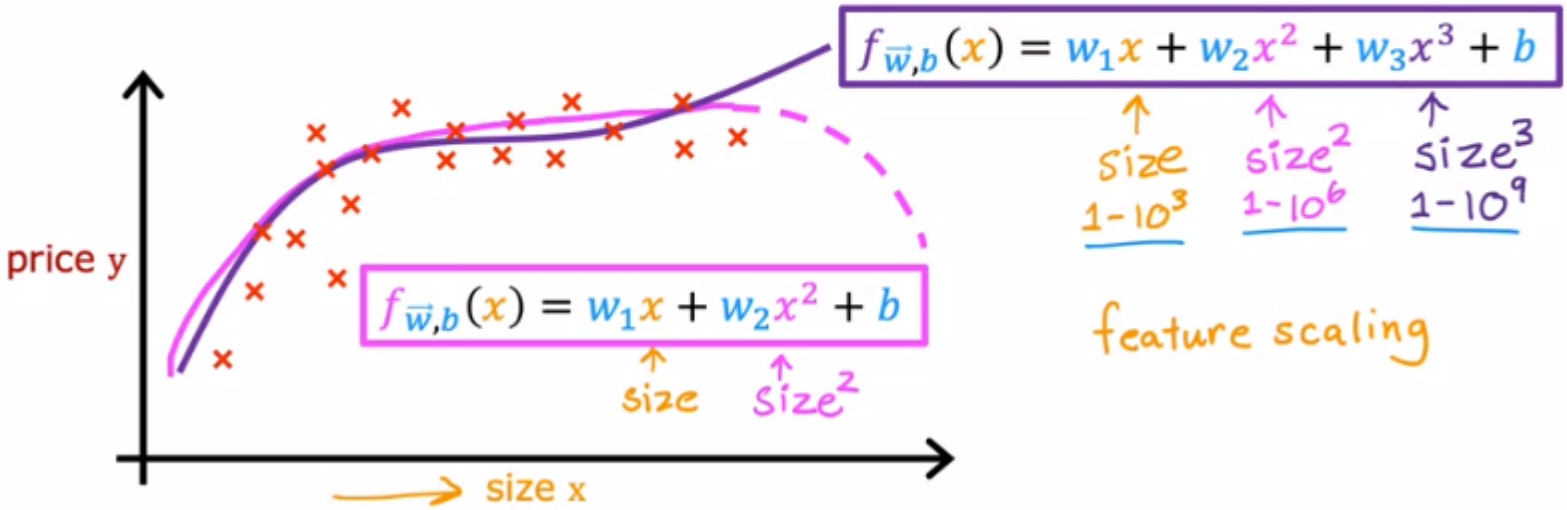
... 0.001 0.003 0.01 0.03 0.1 0.3 1 ...

3X ≈3X 3X ≈3X 3X ≈3X





Polynomial regression



Feature scaling

aim for about $-1 \le x_j \le 1$ for each feature x_j

$-3 \le x_j \le 3$

$-0.3 \le x_j \le 0.3$

} acceptable ranges

$0 \le x_1 \le 3$ okay, no rescaling

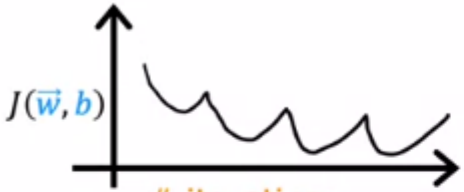
$-2 \le x_2 \le 0.5$ okay, no rescaling

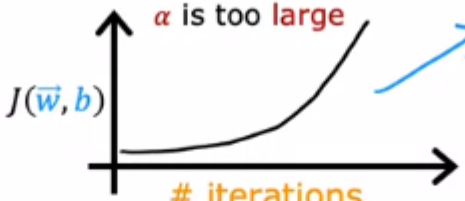
$-100 \le x_3 \le 100$ too large → rescale

$-0.001 \le x_4 \le 0.001$ too small → rescale

$98.6 \le x_5 \le 105$ too large → rescale

Identify problem with gradient descent





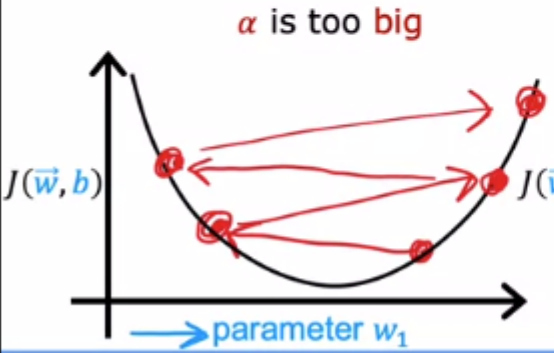
or learning rate is too large

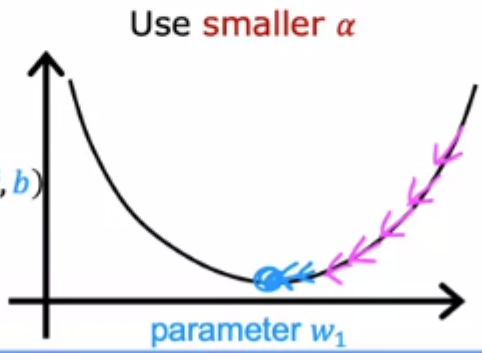
$w_1 = w_1 + \alpha d_1$

use a minus sign

$w_1 = w_1 - \alpha d_1$

Adjust learning rate

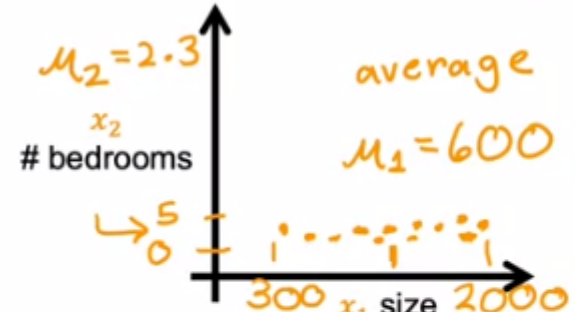




With a small enough α, $J(\vec{w}, b)$ should decrease on every iteration

If α is too small, gradient descent takes a lot more iterations to converge

Mean normalization



$$300 \le x_1 \le 2000$$

$$x_1 = \frac{x_1 - \mu_1}{2000 - 300}$$

max-min

$$-0.18 \le x_1 \le 0.82$$

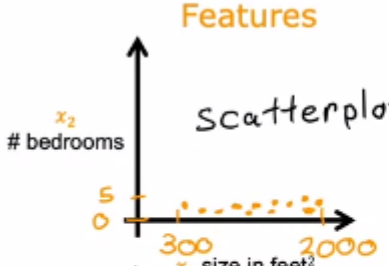
$$0 \le x_2 \le 5$$

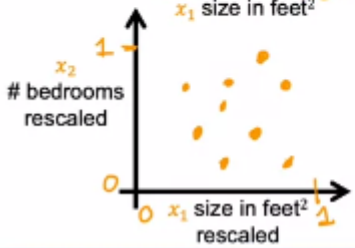
$$x_2 = \frac{x_2 - \mu_2}{5 - 0}$$

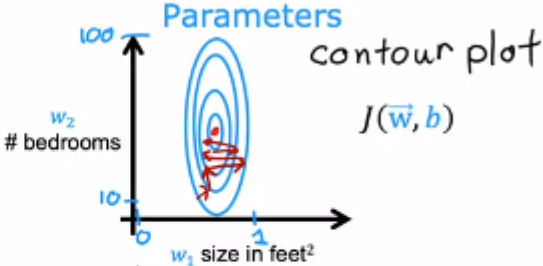
max-min

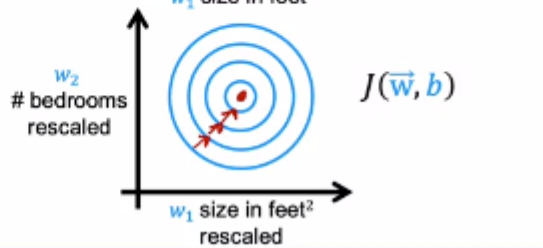
$$-0.46 \le x_2 \le 0.54$$

Feature size and gradient descent

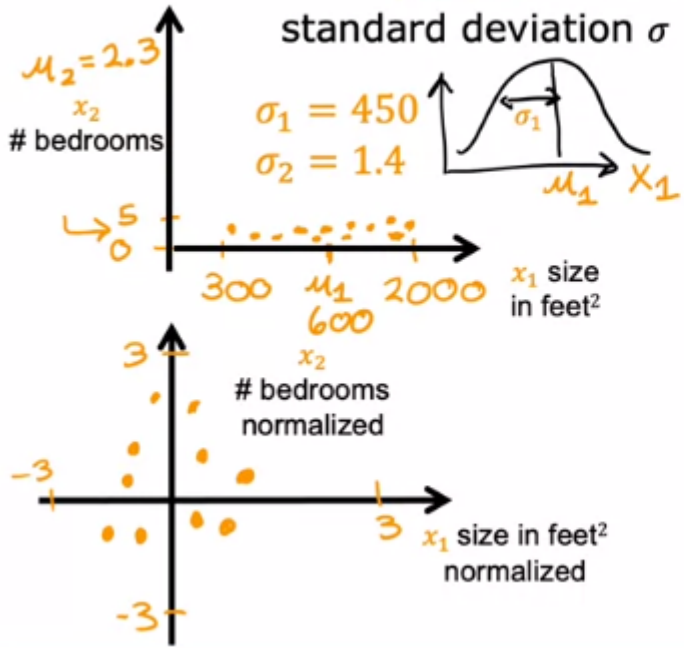








Z-score normalization



$$300 \le x_1 \le 2000 \qquad 0 \le x_2 \le 5$$

$$x_1 = \frac{x_1 - \mu_1}{\sigma_1} \qquad x_2 = \frac{x_2 - \mu_2}{\sigma_2}$$

$$-0.67 \le x_1 \le 3.1 \qquad -1.6 \le x_2 \le 1.9$$