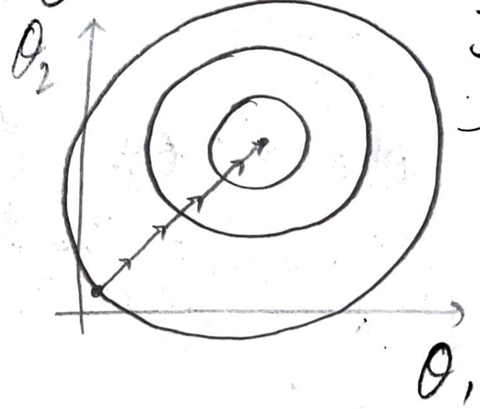
But scaling the feature properly by multiplying or dividing the inputs by an appropriate

$x_1 = \frac{\text{size (feet}^2)}{2000}$; $x_2 = \frac{\text{\# bedrooms}}{5}$ $\quad 0 \le x_1 \le 1$, $\; 0 \le x_2 \le 1$



$J(\theta)$

∴ contour of func^n is less skewed ∴, reaching to minima is easier

Get every feature into approx. a $\boxed{-1 \le x_i \le 1}$ range. where, $x_0 = 1$

Eg) $0 \le x_1 \le 3$ ✓ ; $-2 \le x_2 \le 0.5$ ✓ ; $-100 \le x_3 \le 100$ ✗ ; $-0.0001 \le x_4 \le 0.0001$ ✗

Mean Normalization

Replace x_i with $x_i - \mu_i$ to make features have approximately zero mean (Don't apply $x_0 = 1$)

Eg) $x_1 = \frac{\text{size} - 1000}{2000}$; $x_2 = \frac{\text{\#bedrooms} - 2}{5}$

$\underbrace{1000}$ avg. size ; $\underbrace{2}$ avg # bedroom

$-0.5 \le x_1 \le 0.5$; $-0.5 \le x_2 \le 0.5$

generally:

$x_1 \rightarrow \frac{x_1 - \mu_1}{s_1}$ → μ_1: avg. value of x_1 in training set

s_1 → range (max − min)
→ std deviation of x_1

IIIly,

$x_2 \rightarrow \frac{x_2 - \mu_2}{s_2}$