

یادگیری ماشین با پایتون

فرزاد حسینیان

دانشگاه محقق اردبیلی

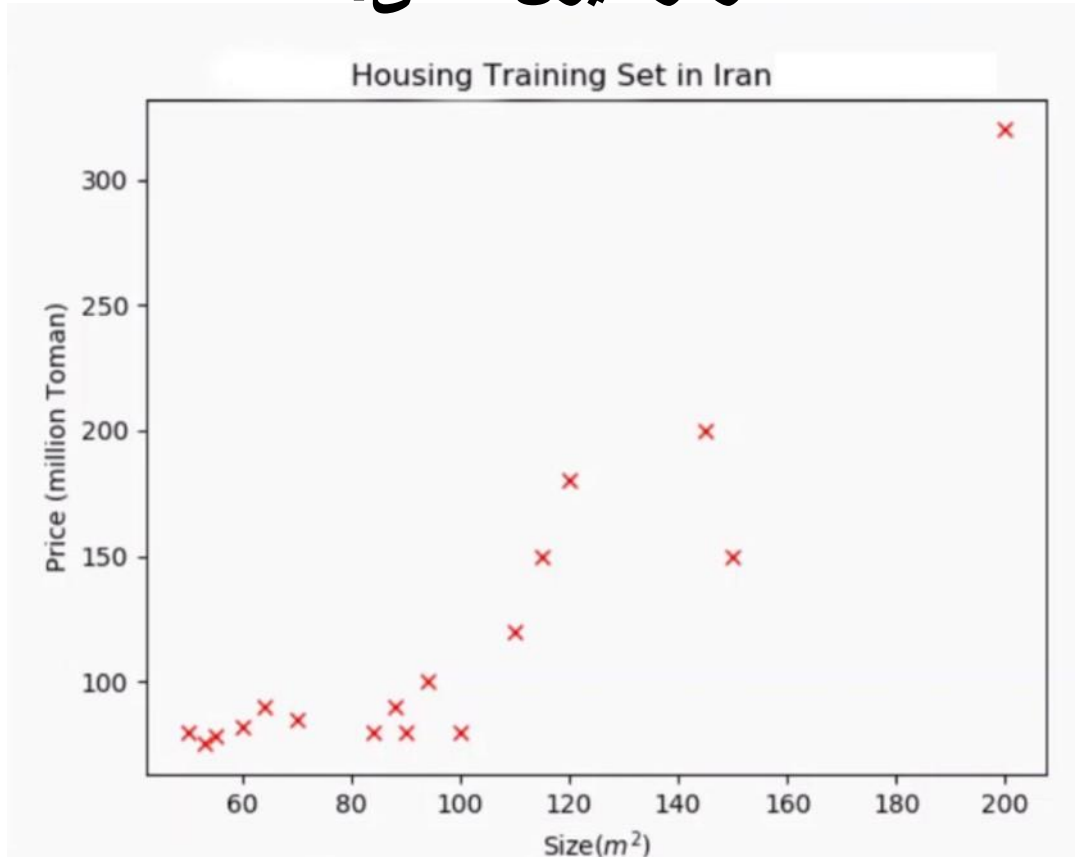
دانشکده علوم کامپیوتر

آبان 98

جلسه یکم

رگرسیون خطی

رگرسیون خطی:



$$h_{\theta}(x) = \theta_0 + \theta_1 x$$

Parameters:

$$\theta_0, \theta_1$$

هدف از رگرسیون: پیدا کردن مدلی که به خوبی رفتار داده ها را شرح دهد

بررسی رگرسیون از دید آماری

مدل ارتباطی: رابطه خطی بین متغیر مستقل و متغیر وابسته

$$Y/x = a + bx + E$$

$$E = Y/x - a - bx$$

• روش کمترین مربع خطا - Least Square Error

- برآورد پارامترها براساس کمینه کردن مربع خطا

$$SSE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n [y_i - (a + bx_i)]^2 = A(a, b)$$

$$\hat{b} = \frac{s_{xy}}{s_x^2} = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

مرور - مشتق جزئی

$$f(x) = x$$

$$\frac{df}{dx} = 1$$

$$f(x, y) = x + y$$

$$\frac{\partial f}{\partial x} = 1$$

$$\frac{\partial f}{\partial y} = 1$$

$$f(x, y) = xy$$

$$\frac{\partial f}{\partial x} = y$$

$$\frac{\partial f}{\partial y} = x$$

تابع هزینه – Cost Function

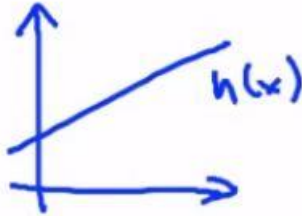
معیار یادگیری برای میزان مناسب بودن
مدل فیت شده به دیتا ها

Hypothesis:

$$\underline{h_{\theta}(x) = \theta_0 + \theta_1 x}$$

Parameters:

$$\underline{\theta_0, \theta_1}$$



میانگین مجذور خطا مدل برای همه داده ها

Cost Function:

$$\rightarrow J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

Goal: minimize $J(\theta_0, \theta_1)$
 $\nearrow \theta_0, \theta_1$

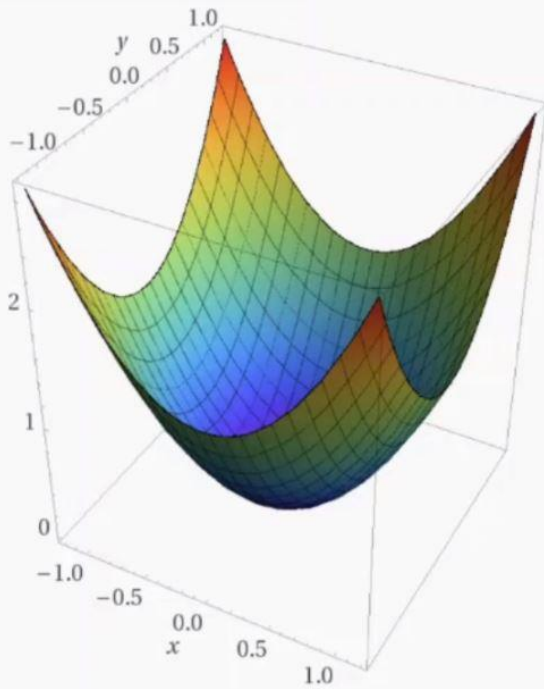
ایده کمینه کردن خطاها

هدف $\leq$ کمینه کردن مجذور خطاها

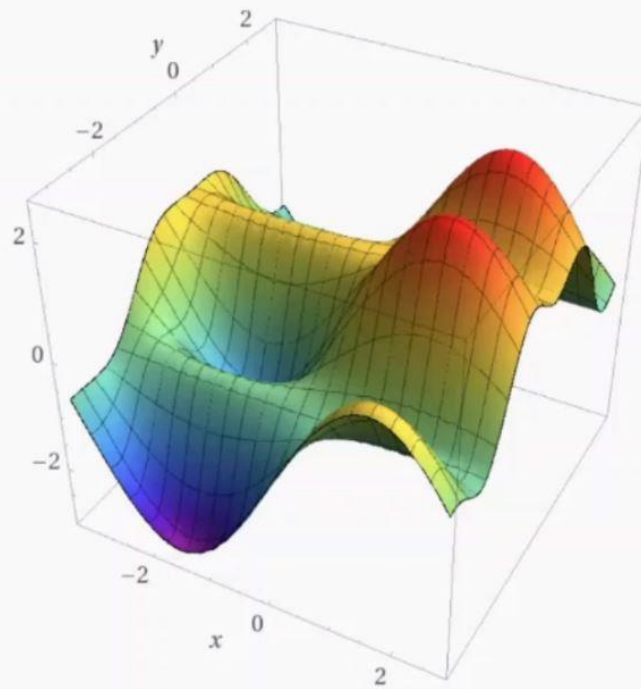
ایده:

1. با یکسری از تتا های تصادفی شروع کنیم
2. با تغییر دادن این تتاها امید به کاهش میزان خطا داشته باشیم

آیا تضمینی برای رسیدن به کمینه خطا هست؟



convex



Non-convex

گرادیان کاهش

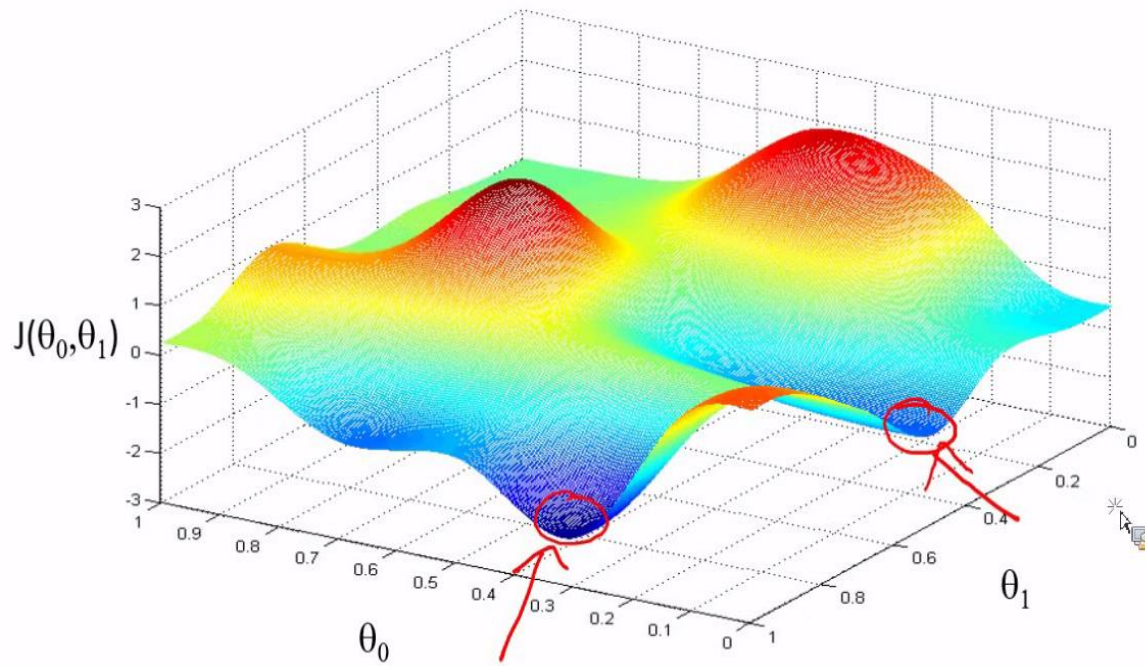
Gradient descent algorithm

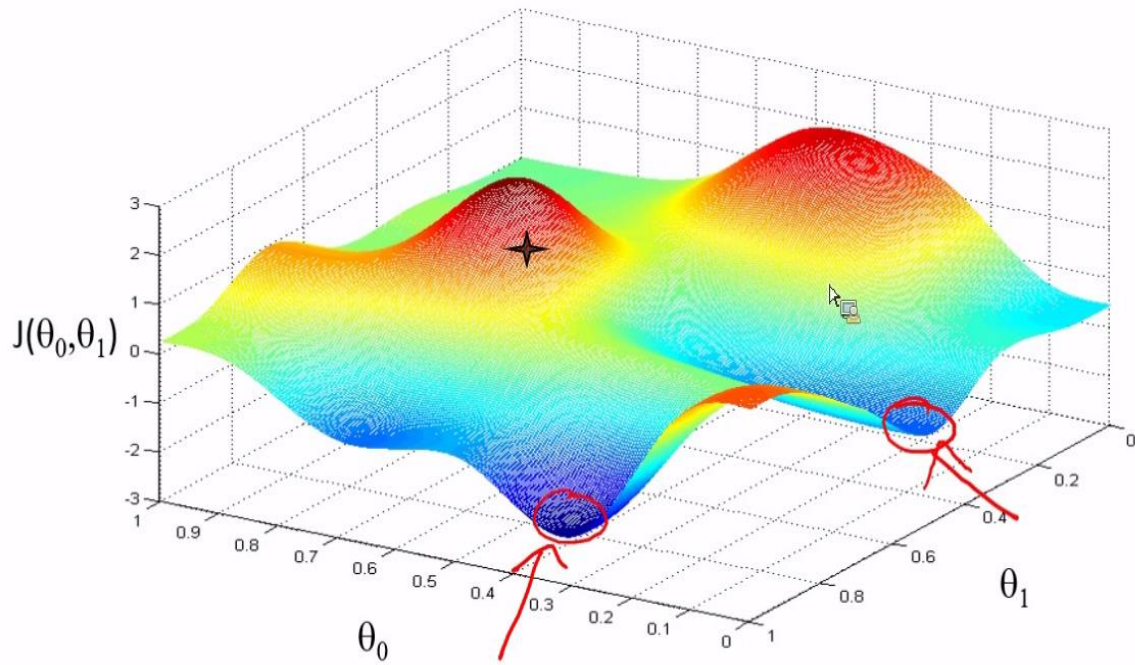
repeat until convergence {
 $\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$
 (for $j = 1$ and $j = 0$)
}

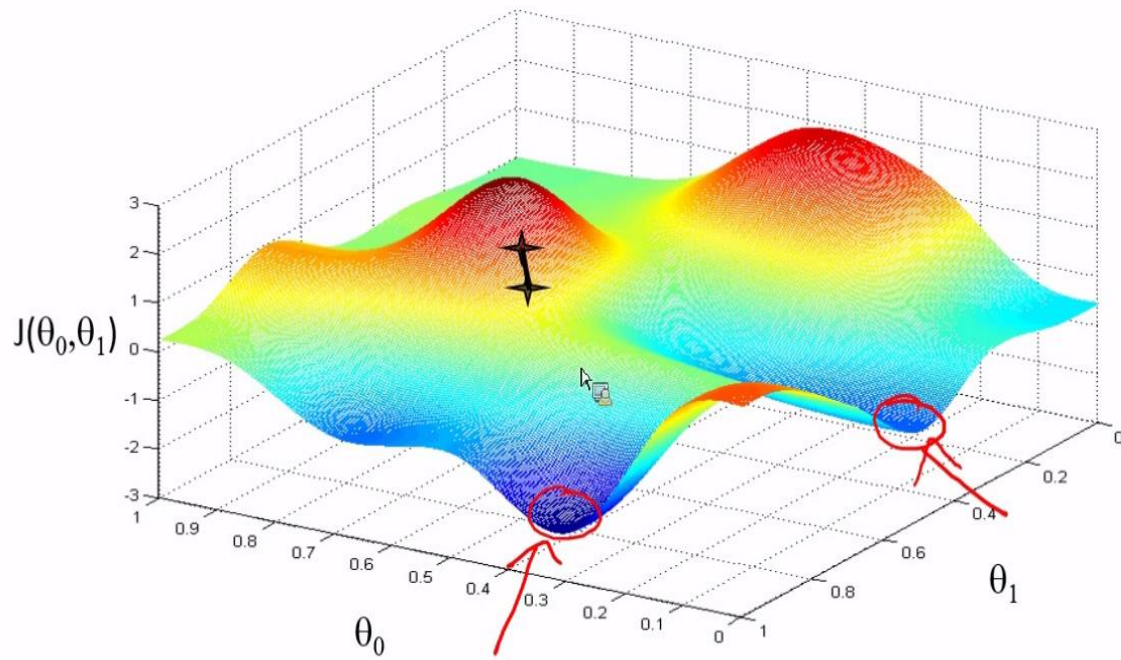
Linear Regression Model

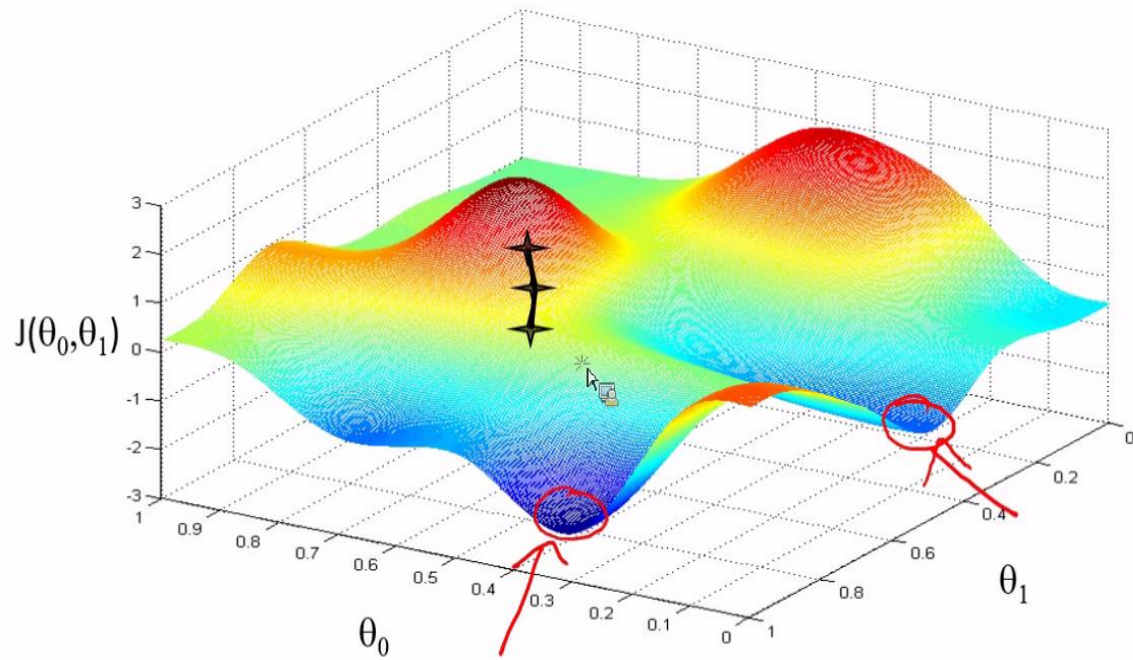
$$h_{\theta}(x) = \theta_0 + \theta_1 x$$

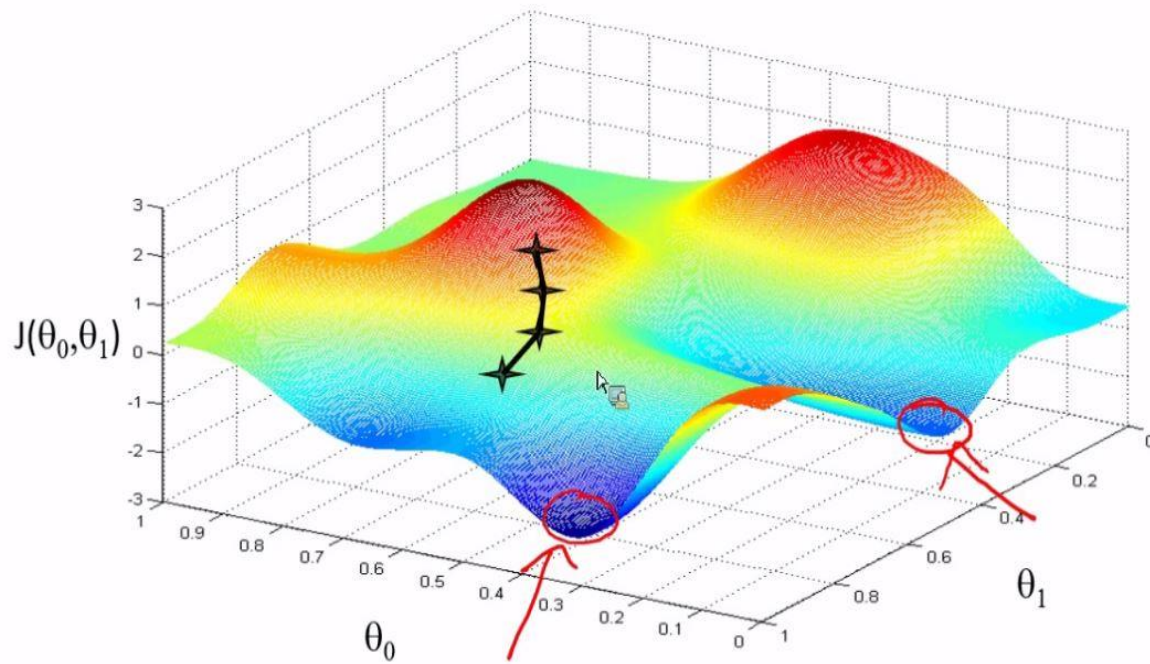
$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

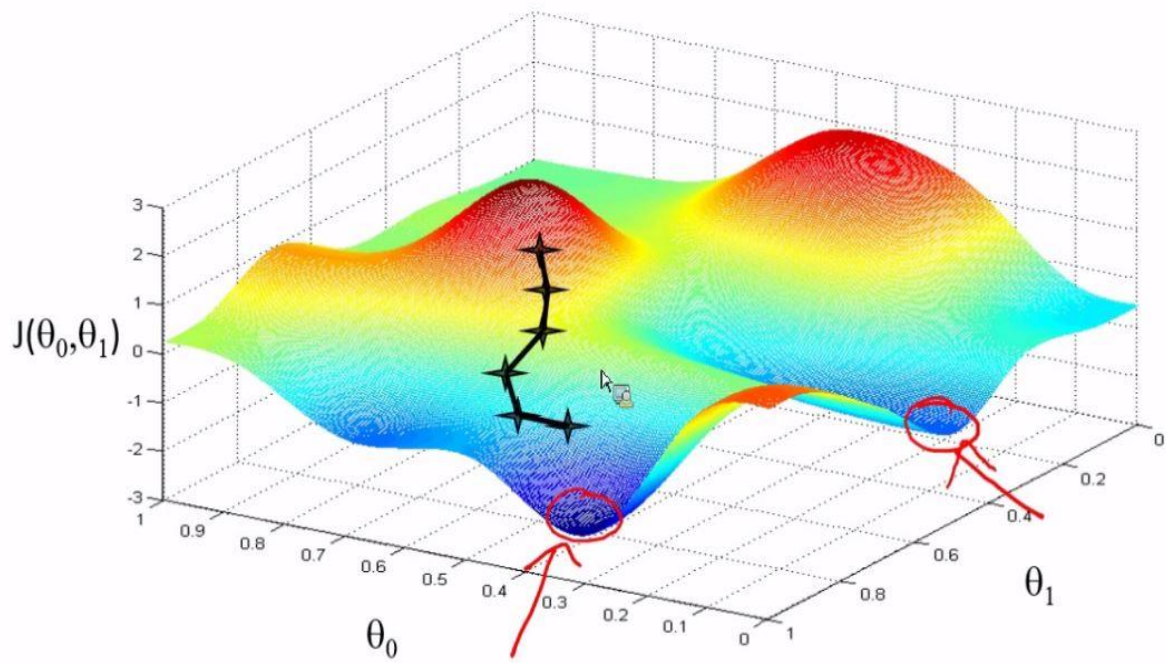


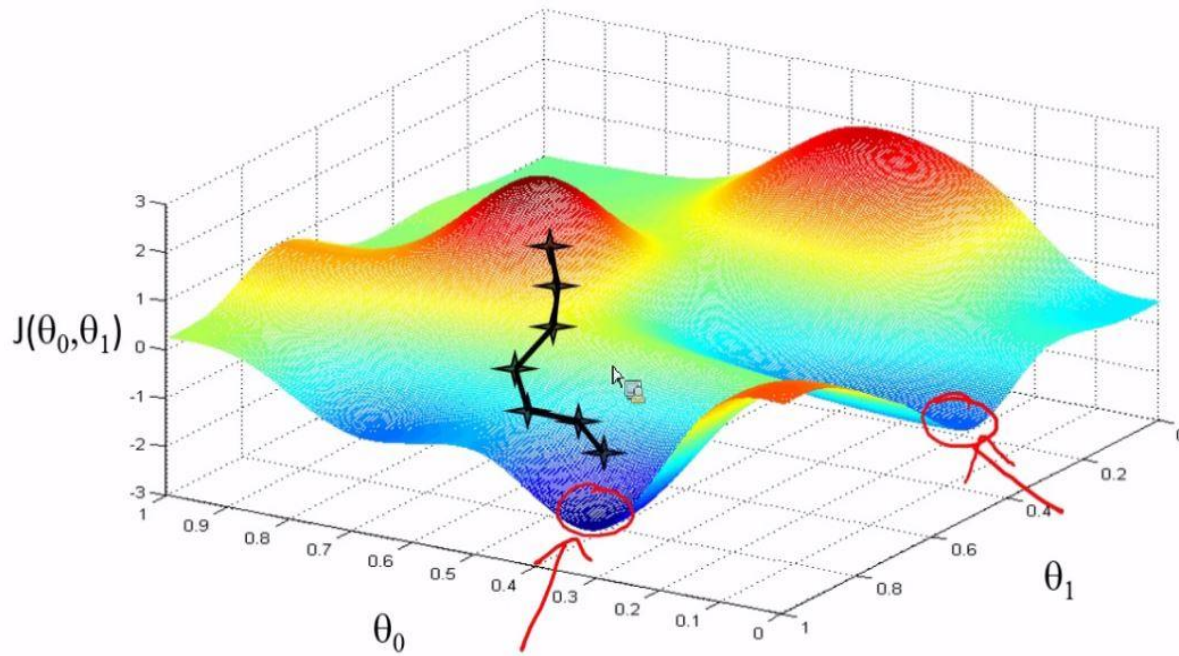


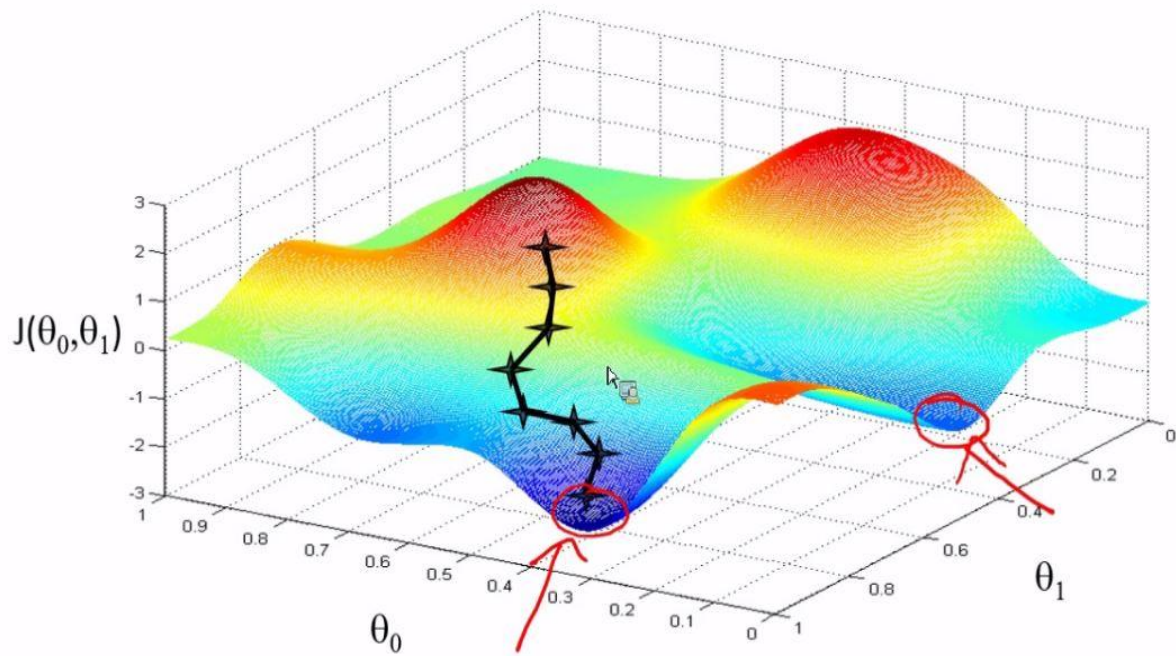












گرادیان تابع هزینه

Gradient descent algorithm

repeat until convergence {
 $\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$
 (for $j = 1$ and $j = 0$)
}

Linear Regression Model

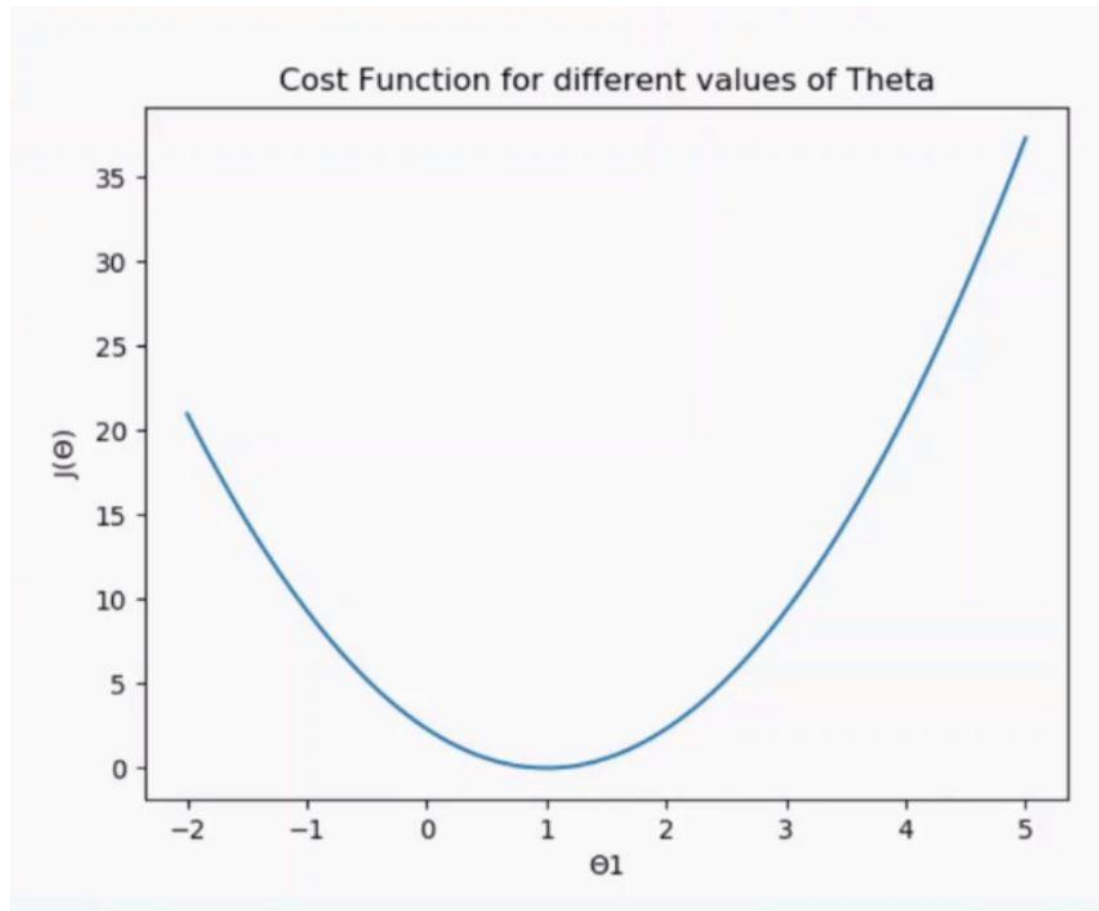
$$h_{\theta}(x) = \theta_0 + \theta_1 x$$

$$J(\theta_0, \theta_1) = \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2$$

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})$$

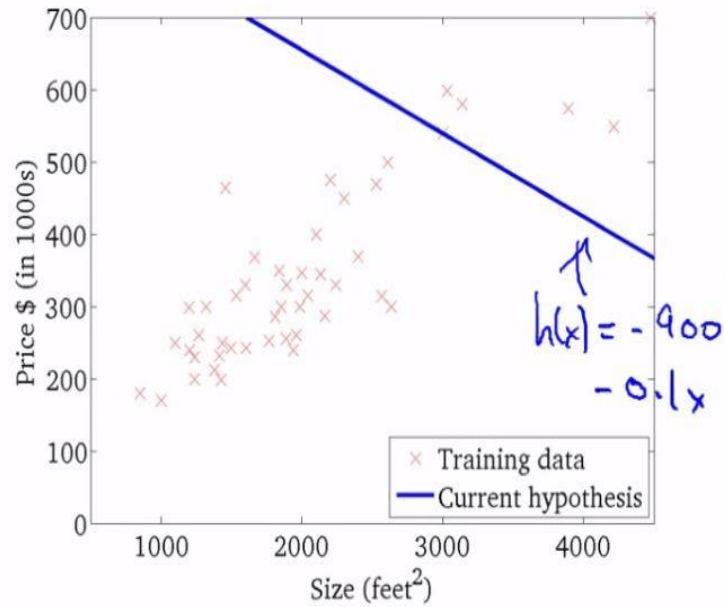
$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x^{(i)}$$

نرخ یادگیری: اندازه قدم – مقداری که تنها را تغییر میدهیم



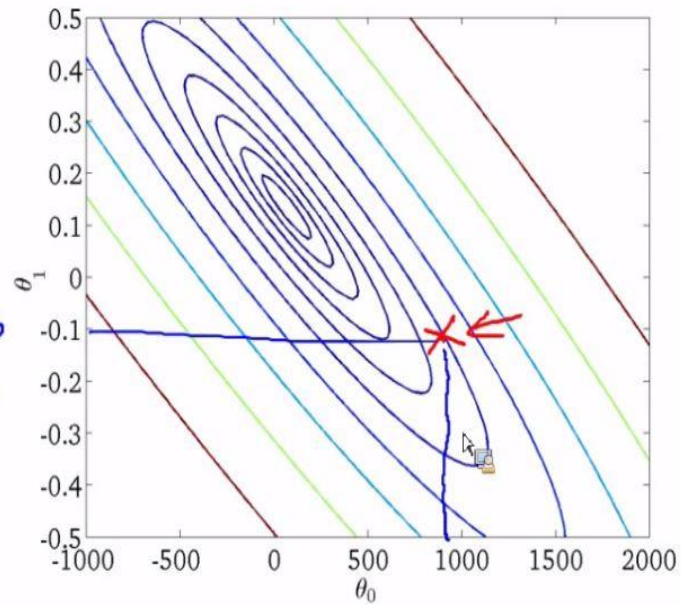
$$h_{\theta}(x)$$

(for fixed θ_0, θ_1 , this is a function of x)



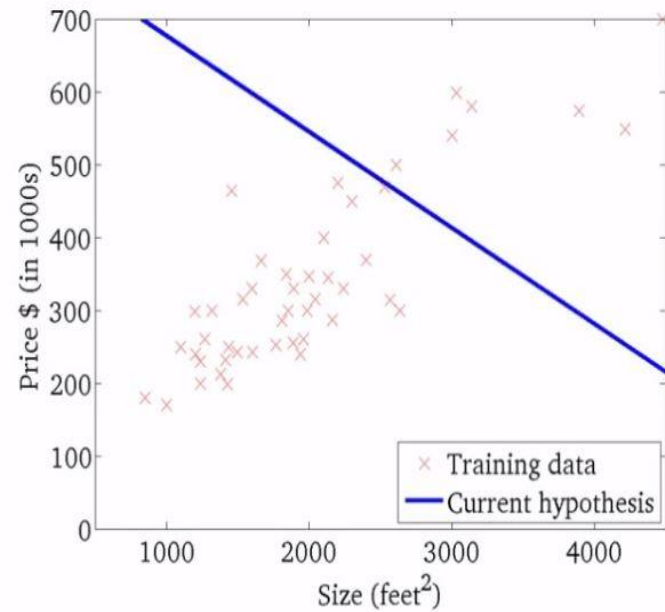
$$J(\theta_0, \theta_1)$$

(function of the parameters θ_0, θ_1)



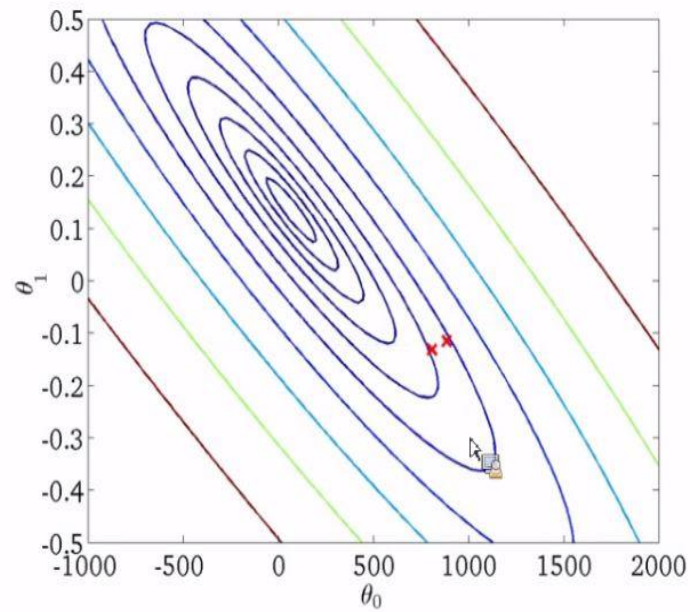
$$h_{\theta}(x)$$

(for fixed θ_0, θ_1 , this is a function of x)



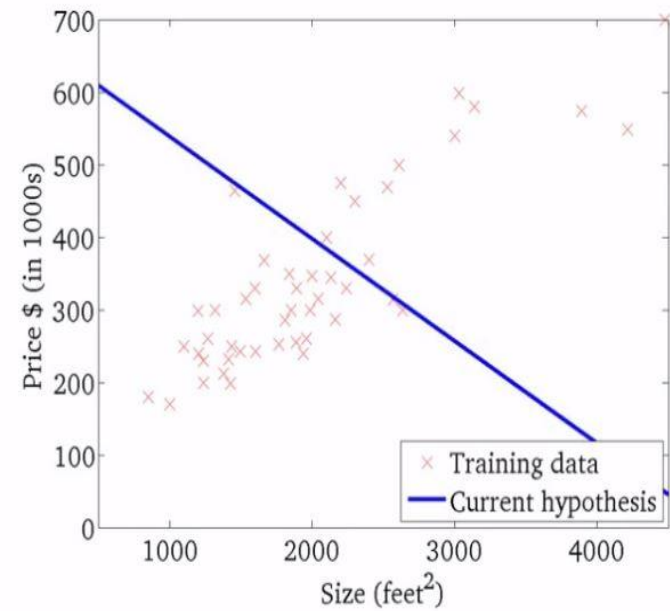
$$J(\theta_0, \theta_1)$$

(function of the parameters θ_0, θ_1)



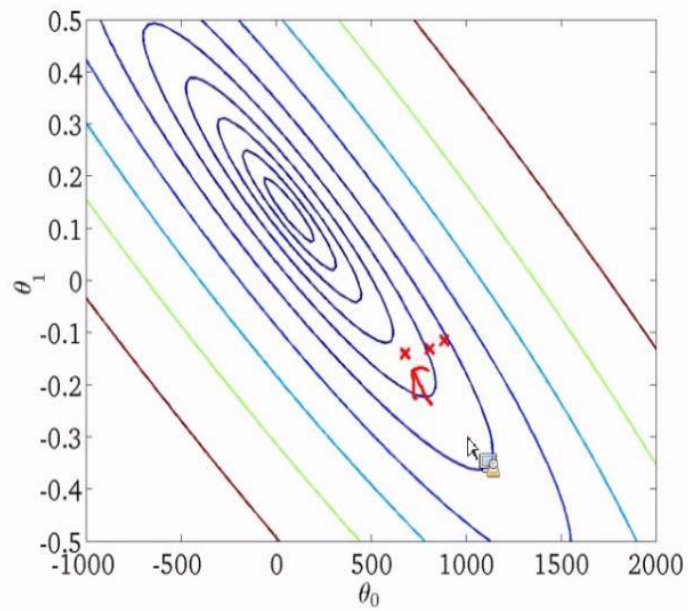
$$h_{\theta}(x)$$

(for fixed θ_0, θ_1 , this is a function of x)



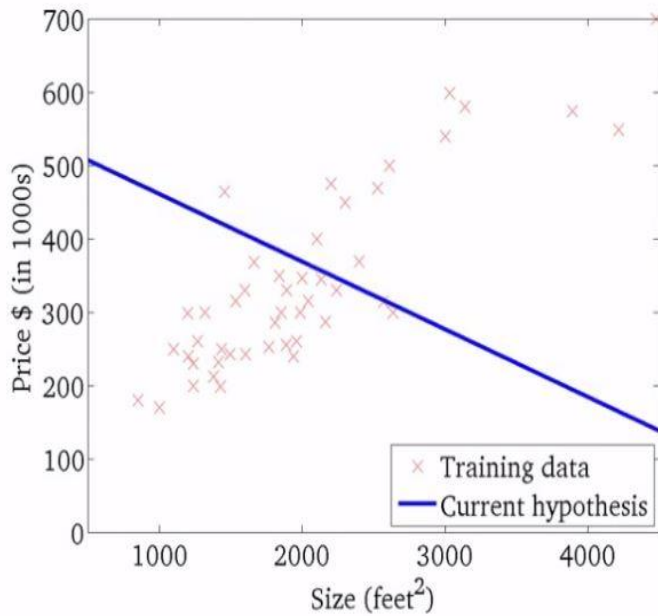
$$J(\theta_0, \theta_1)$$

(function of the parameters θ_0, θ_1)



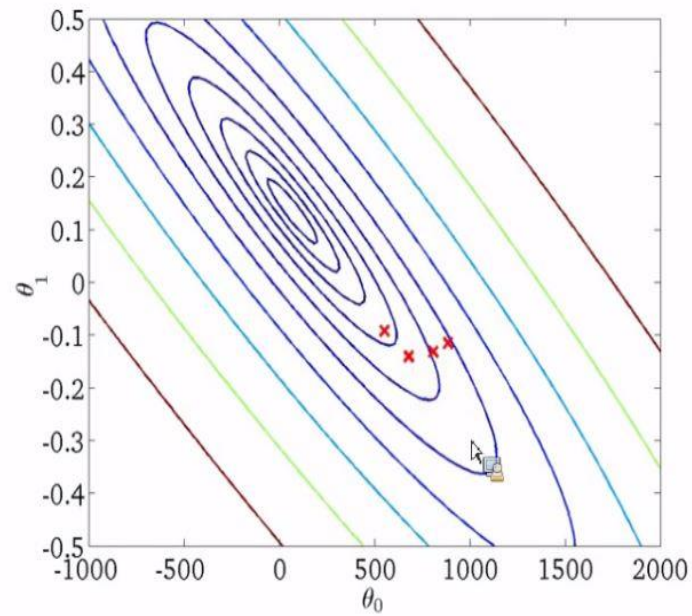
$$h_{\theta}(x)$$

(for fixed θ_0, θ_1 , this is a function of x)



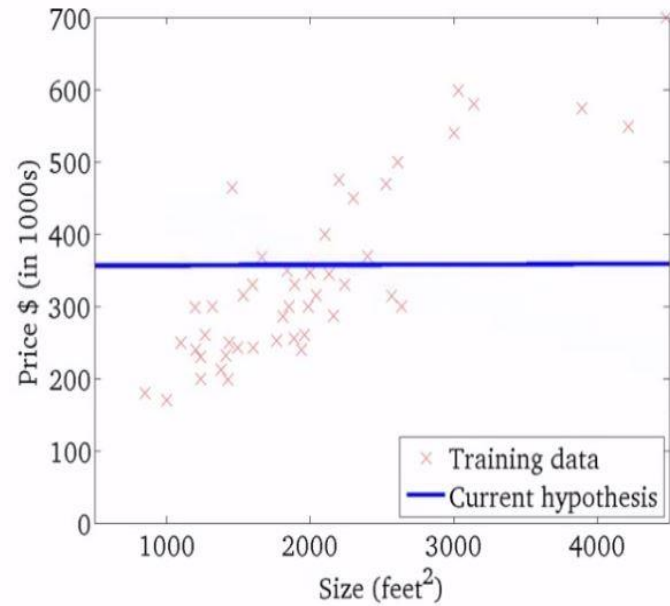
$$J(\theta_0, \theta_1)$$

(function of the parameters θ_0, θ_1)



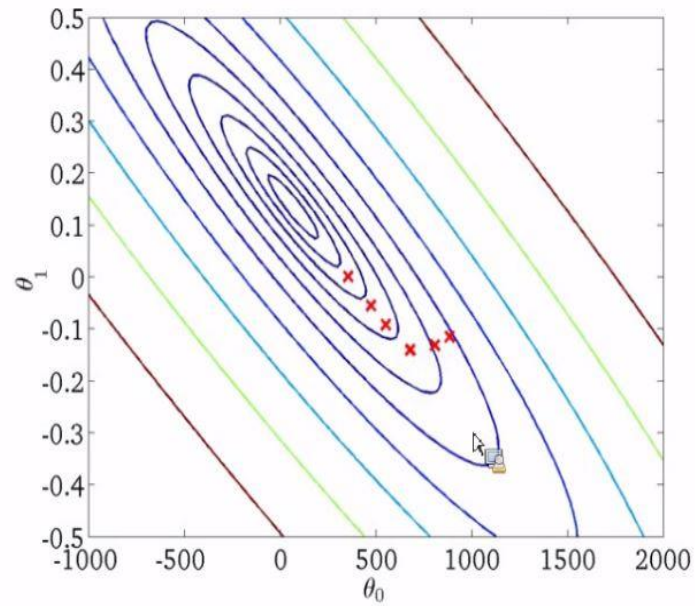
$$h_{\theta}(x)$$

(for fixed θ_0, θ_1 , this is a function of x)



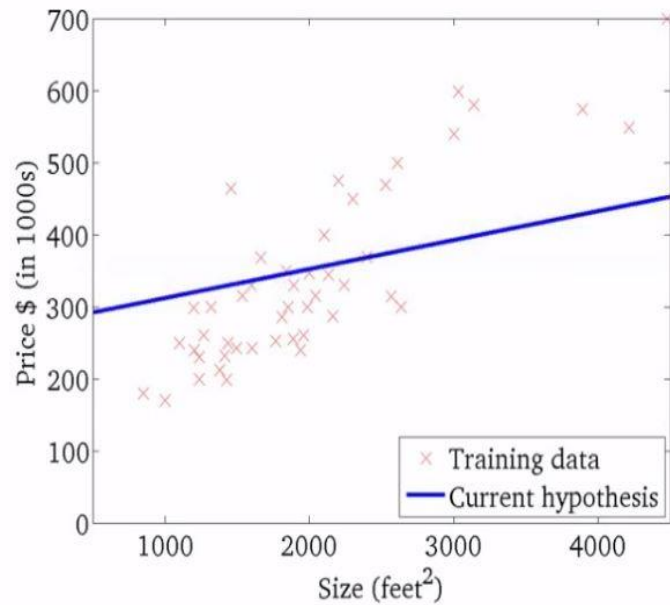
$$J(\theta_0, \theta_1)$$

(function of the parameters θ_0, θ_1)



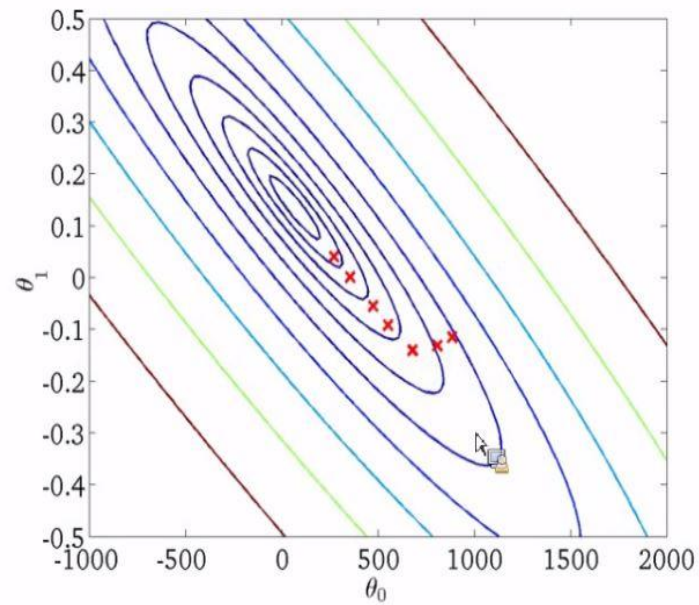
$$h_{\theta}(x)$$

(for fixed θ_0, θ_1 , this is a function of x)



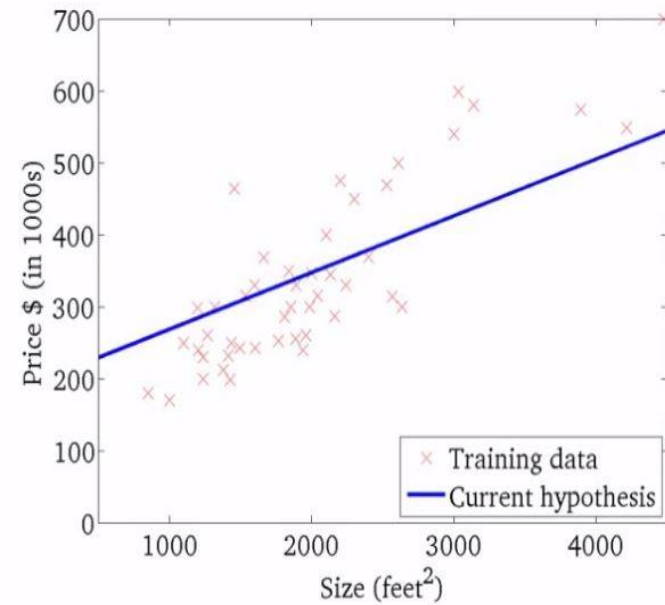
$$J(\theta_0, \theta_1)$$

(function of the parameters θ_0, θ_1)



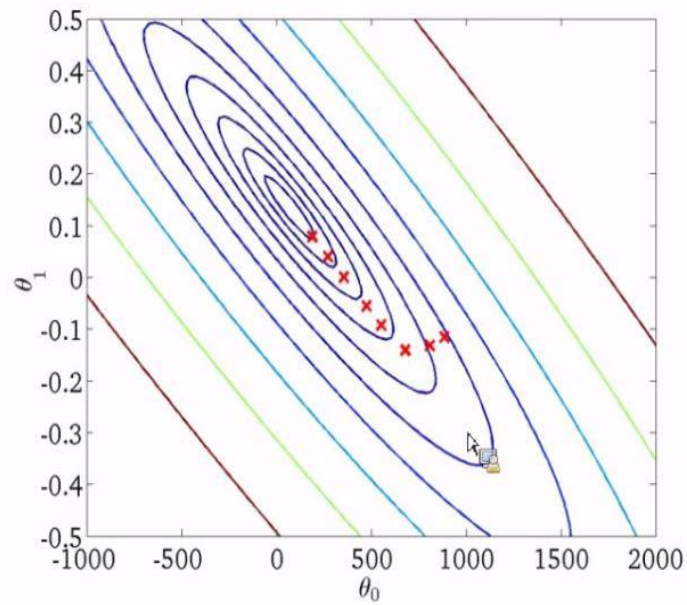
$$h_{\theta}(x)$$

(for fixed θ_0, θ_1 , this is a function of x)



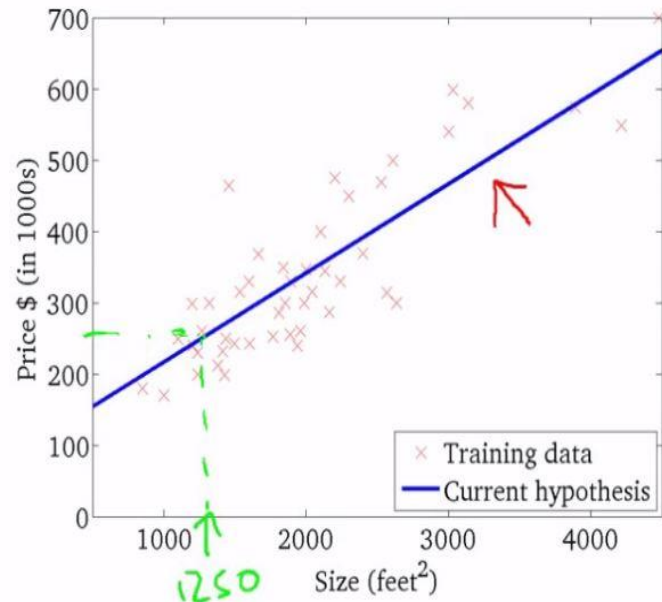
$$J(\theta_0, \theta_1)$$

(function of the parameters θ_0, θ_1)



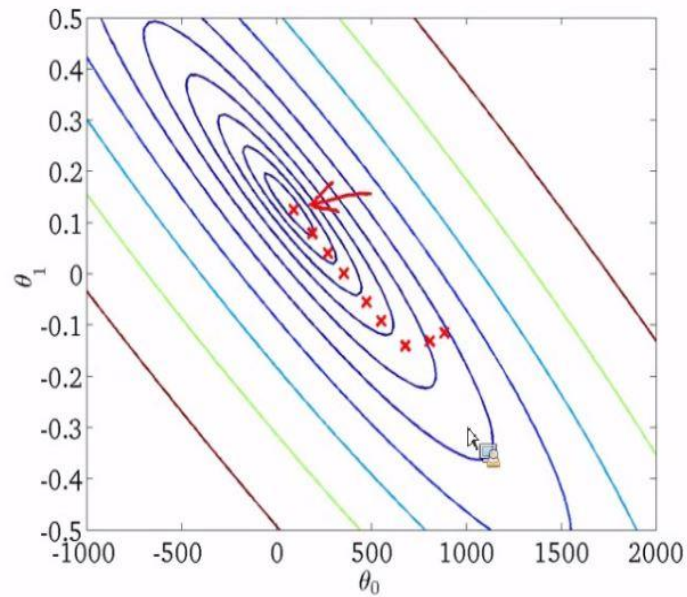
$$h_{\theta}(x)$$

(for fixed θ_0, θ_1 , this is a function of x)



$$J(\theta_0, \theta_1)$$

(function of the parameters θ_0, θ_1)



تمرین

کد و تحلیل به روش SSE

مثال: نمرات مربوط به درس زبان خارجی برای ۱۰ نفر که تعداد سال‌های اقامتشان در یک کشور خارجی ثبت شده، در جدول آمده‌است. برآورد مدل رگرسیون خطی به صورت زیر است:

جمع

مدت اقامت	-X	3	4	4	2	5	3	4	5	3	2	35
-Y	نمره زبان خارجی	57	78	72	58	89	63	73	84	75	48	697