#### Intro to Machine Learning (CS436/CS580L)

# Lecture 5 & 6: Linear Regression & Gradient Descent & Polynomial Regression

Xi Peng, Fall 2018

Thanks to Tom Mitchell, Andrew Ng, Ben Taskar, Carlos Guestrin, Eric Xing, Hal Daume III, David Sontag, Jerry Zhu, Tina Eliassi-Rad, and Chao Chen for some slides & teaching material.

## This Class

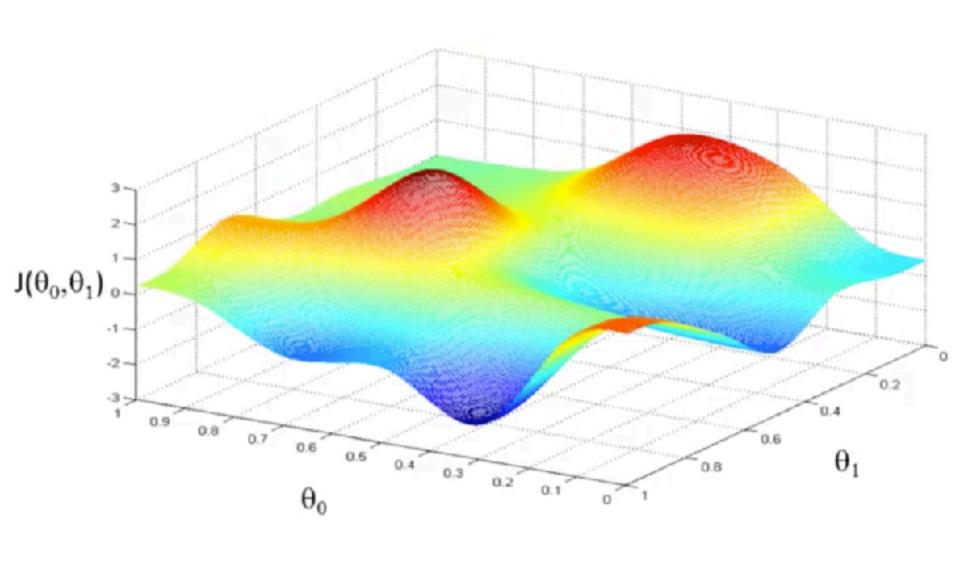
- Linear Regression (One variable)
- Gradient Descend
- Linear Regression (Multiple variables)
- Polynomial Regression

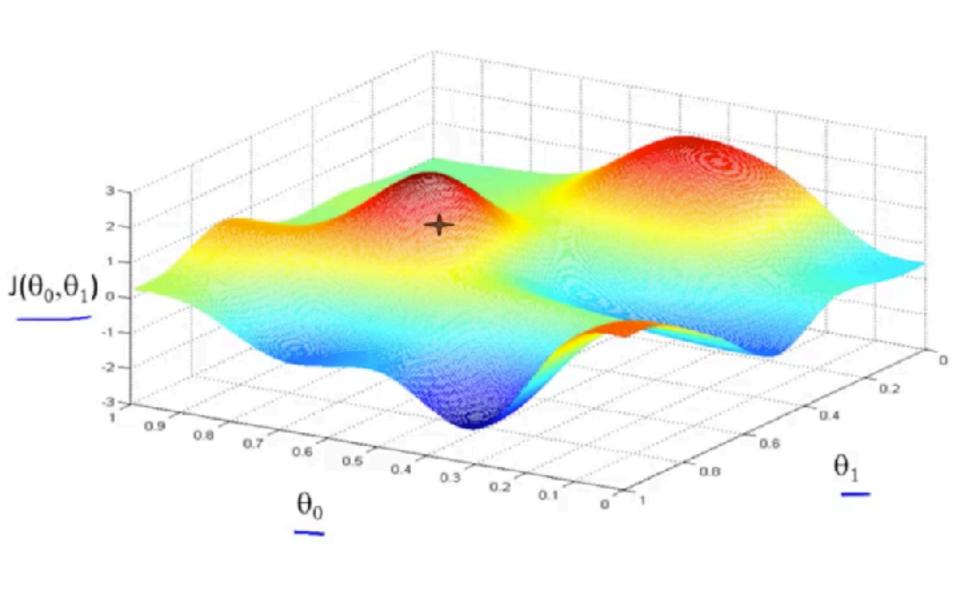
### Median

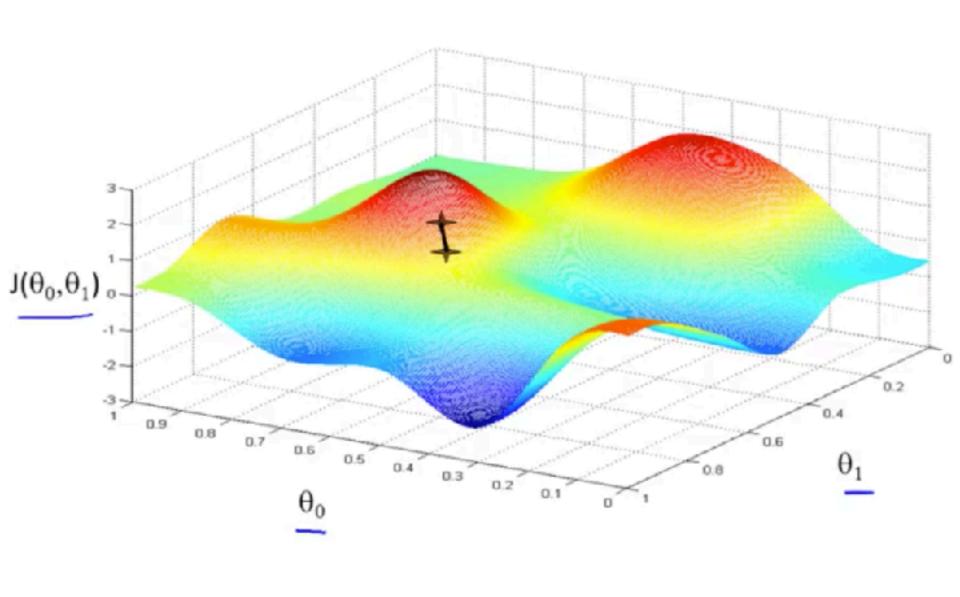
Have some function  $J(\theta_0,\theta_1)$  Want  $\min_{\theta_0,\theta_1} J(\theta_0,\theta_1)$ 

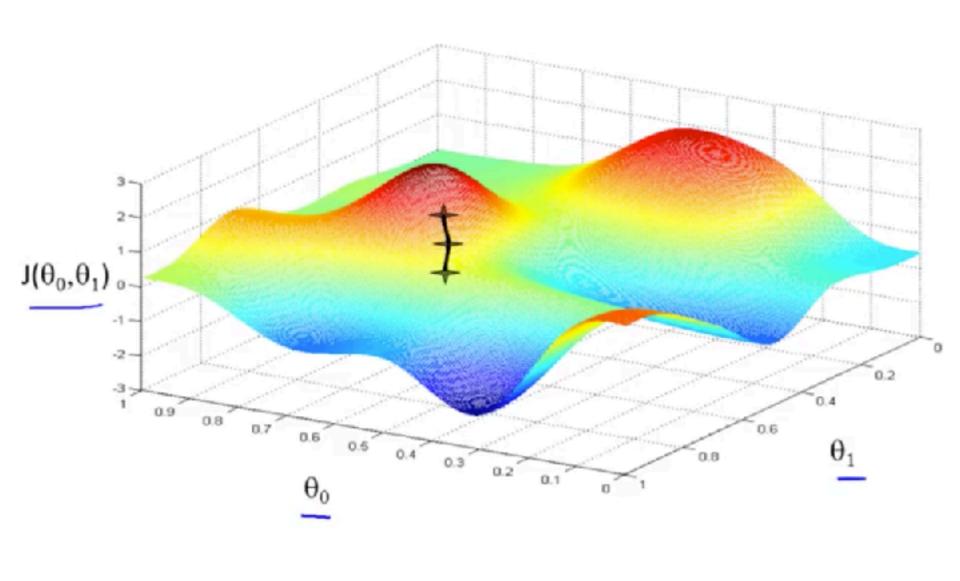
#### Outline:

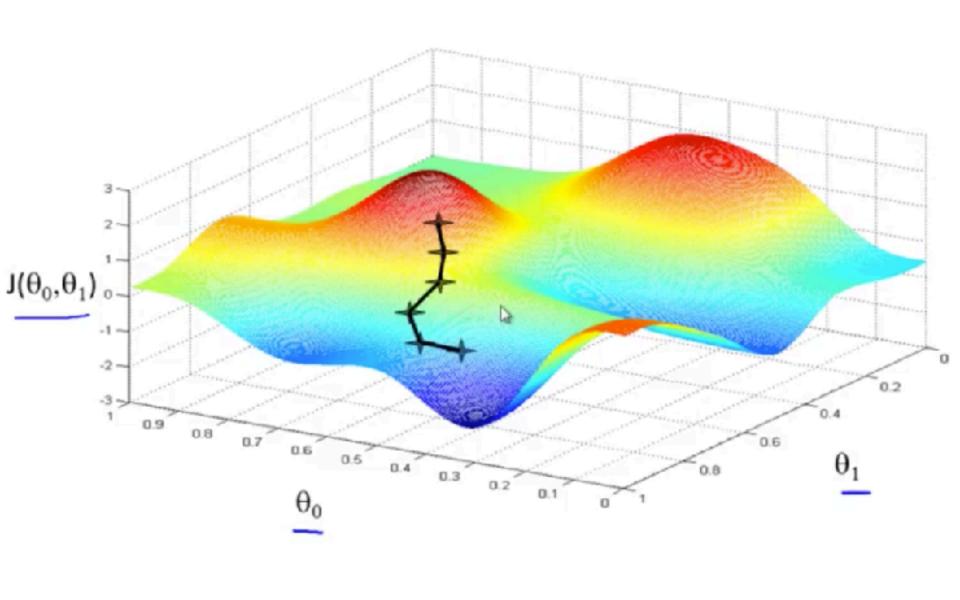
- Start with some  $heta_0, heta_1$
- Keep changing  $heta_0, heta_1$  to reduce  $J( heta_0, heta_1)$  until we hopefully end up at a minimum

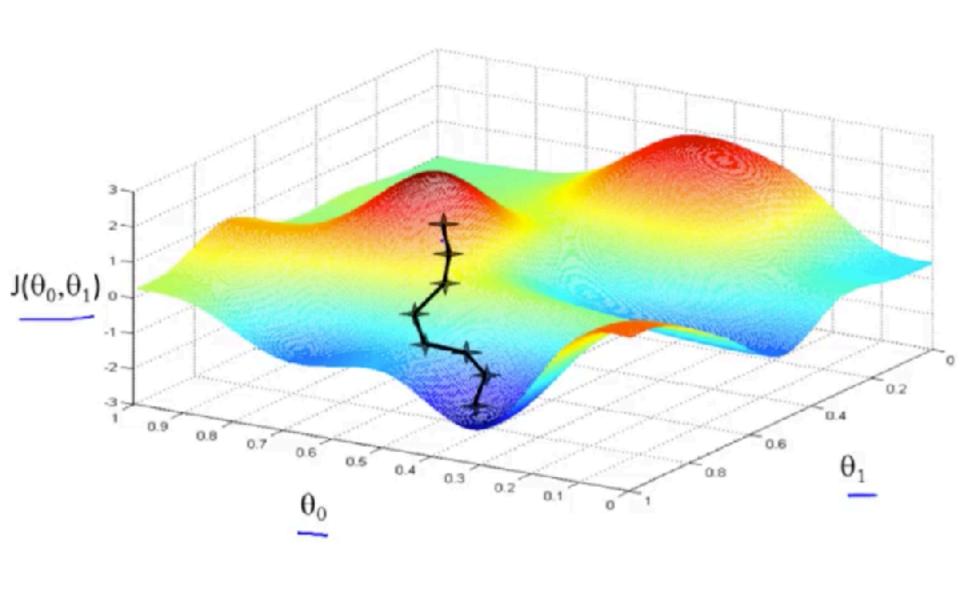


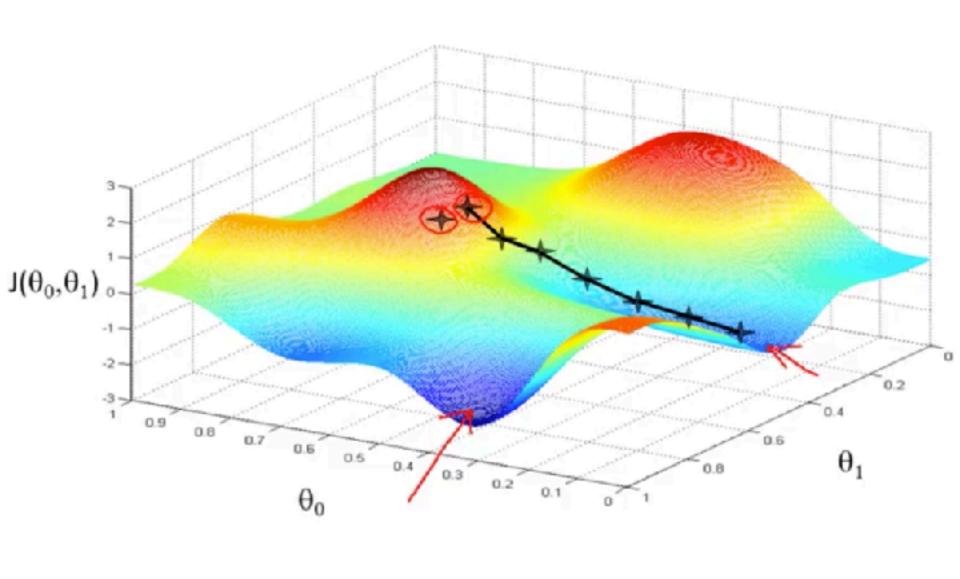






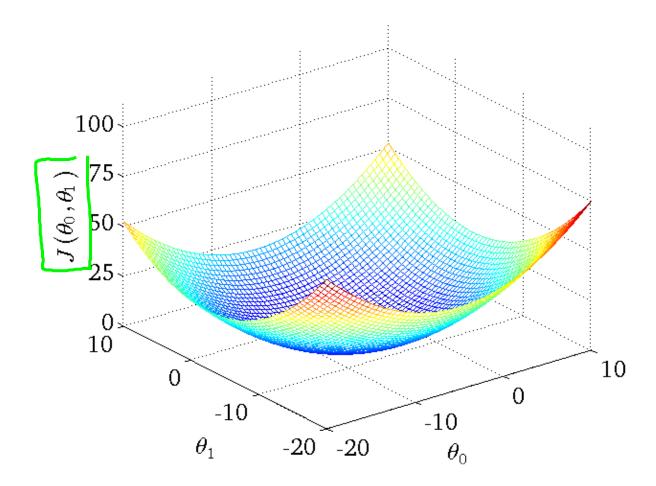






#### Gradient descent algorithm

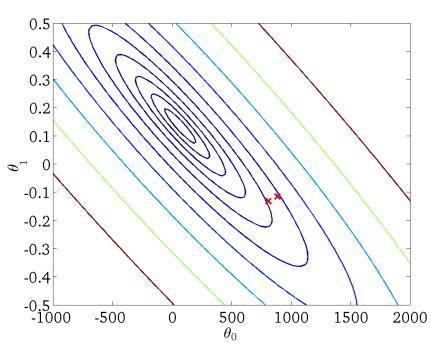
```
repeat until convergence {
\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^m \left( h_{\theta}(x^{(i)}) - y^{(i)} \right)
\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^m \left( h_{\theta}(x^{(i)}) - y^{(i)} \right) \cdot x^{(i)}
}
```



 $h_{\theta}(x)$ (for fixed  $\theta_0, \theta_1$ , this is a function of x) 700 600 Price \$ (in 1000s) 500 200 100 Training data Current hypothesis 0 1000 2000

Size (feet<sup>2</sup>)

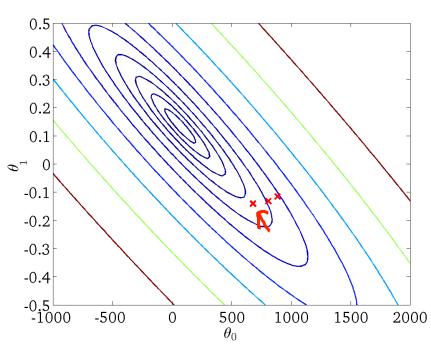
 $J(\theta_0, \theta_1)$ (function of the parameters  $\theta_0, \theta_1$ )



 $h_{\theta}(x)$ (for fixed  $\theta_0, \theta_1$ , this is a function of x) 700 600 500 200 100 Training data Current hypothesis 0 1000 2000 3000 4000

Size (feet<sup>2</sup>)

 $J( heta_0, heta_1)$  (function of the parameters  $heta_0, heta_1$ )



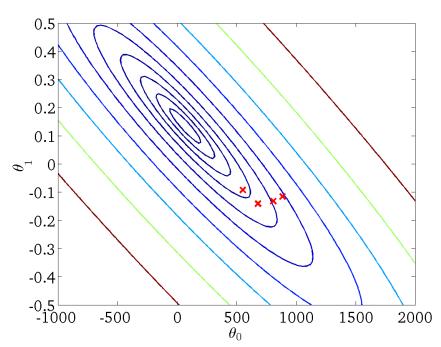
 $h_{\theta}(x)$ (for fixed  $\theta_0, \theta_1$ , this is a function of x) 700 600 Price \$ (in 1000s) 000 \$ 300 000 \$ 500 500 200 100 Training data Current hypothesis 0

3000

Size (feet<sup>2</sup>)

4000

 $J( heta_0, heta_1)$  (function of the parameters  $heta_0, heta_1$ )



 $h_{\theta}(x)$ (for fixed  $\theta_0, \theta_1$ , this is a function of x) 700 600 Price \$ (in 1000s)
000 \$ 300
000 \$ 500 500 200 100 Training data Current hypothesis

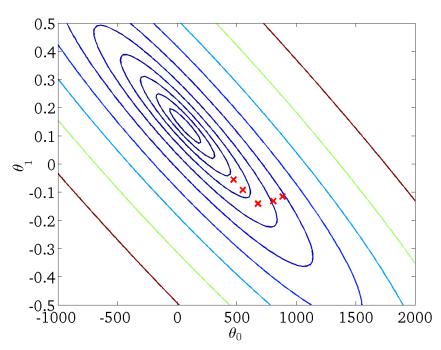
3000

Size (feet<sup>2</sup>)

4000

0

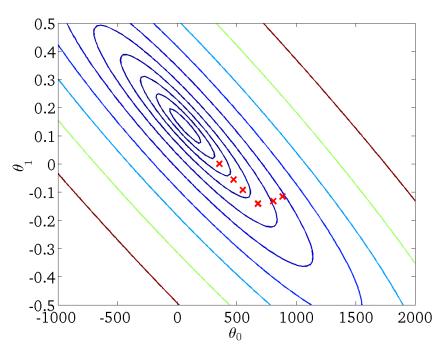
 $J( heta_0, heta_1)$  (function of the parameters  $heta_0, heta_1$ )



 $h_{\theta}(x)$ (for fixed  $\theta_0, \theta_1$ , this is a function of x) 700 600 500 200 100 Training data Current hypothesis 0 1000 2000 3000 4000

Size (feet<sup>2</sup>)

 $J( heta_0, heta_1)$  (function of the parameters  $heta_0, heta_1$ )



 $h_{\theta}(x)$ (for fixed  $\theta_0, \theta_1$ , this is a function of x) 700 600 Price \$ (in 1000s) 000 \$ 000 000 \$ 000 500 200 100 Training data Current hypothesis

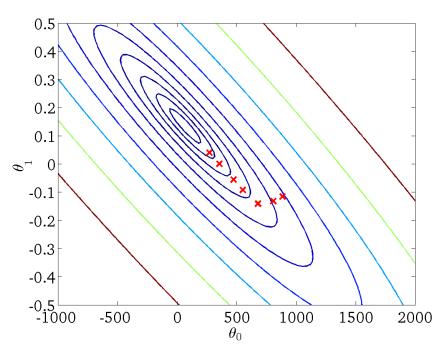
Size (feet<sup>2</sup>)

3000

4000

0

 $J( heta_0, heta_1)$  (function of the parameters  $heta_0, heta_1$ )



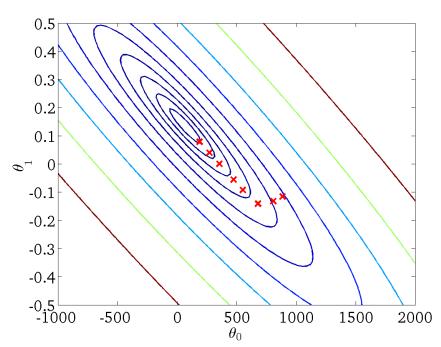
 $h_{\theta}(x)$ (for fixed  $\theta_0$ ,  $\theta_1$ , this is a function of x) 700 600 Price \$ (in 1000s)
000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 500 200 100 Training data Current hypothesis 0

Size (feet<sup>2</sup>)

3000

4000

 $J( heta_0, heta_1)$  (function of the parameters  $heta_0, heta_1$ )



 $h_{\theta}(x)$ (for fixed  $\theta_0, \theta_1$ , this is a function of x) 700 600 Price \$ (in 1000s)
000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 000 \$ 500 100 Training data Current hypothesis 0

Size (feet<sup>2</sup>)

3000

4000

1000

 $J( heta_0, heta_1)$  (function of the parameters  $heta_0, heta_1$ )

