Instituto de Computação
UNIVERSIDADE ESTADUAL DE CAMPINAS

Capacitação profissional em tecnologias de Inteligência Artificial

Machine Learning Overview

Prof. Edson Borin
https://www.ic.unicamp.br/~edson
Institute of Computing - UNICAMP

Linear Regression Example

Linear Regression Example

Goals:

- Introduce key concepts of machine learning by discussing how to train a simple linear regression model to solve a simple task.
- Overview of what ML models look like, their parameters, cost functions, and model fit.

Linear Regression Example

Problem: House price prediction

- Estimate the price of a house based on their size.

70.000 USD

160.000 USD

Linear Regression Example

Problem: House price prediction

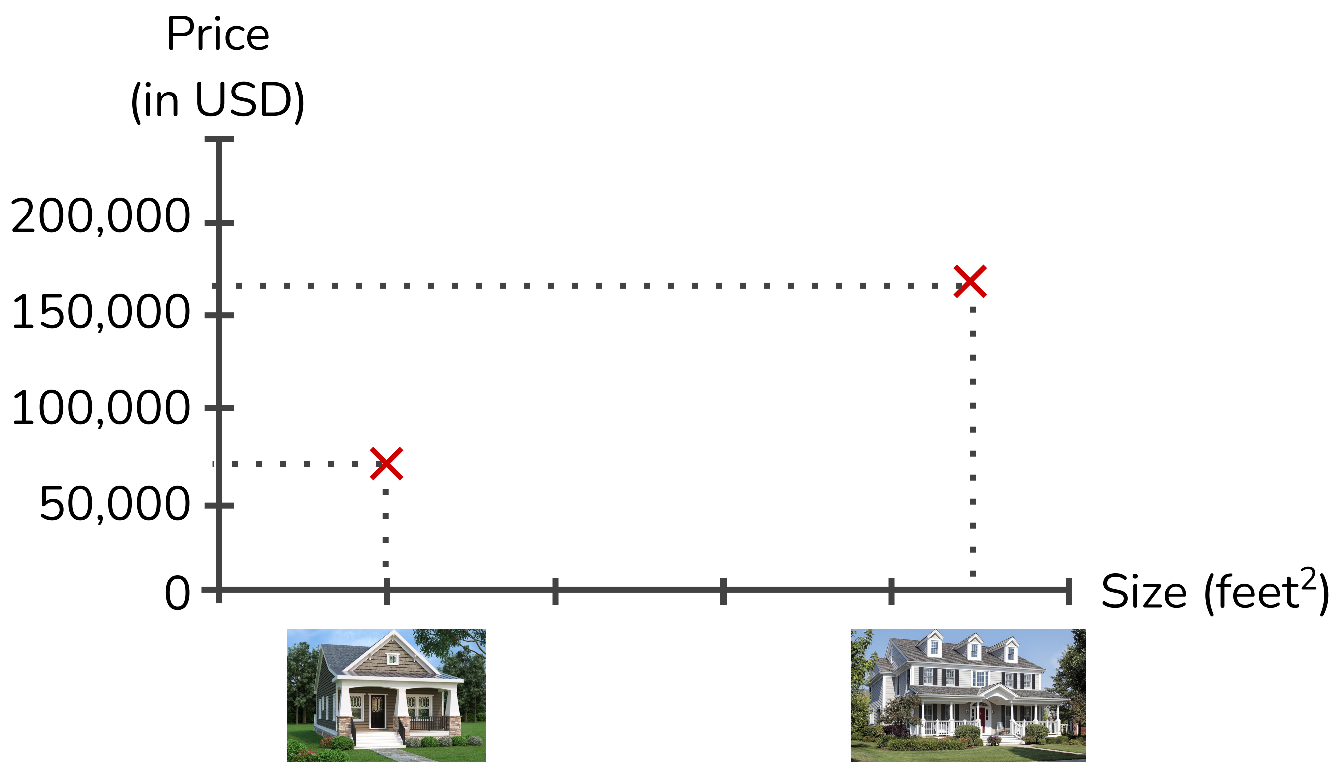
- Estimate the price of a house based on their size.

??? USD

Linear Regression Example

Problem: House price prediction

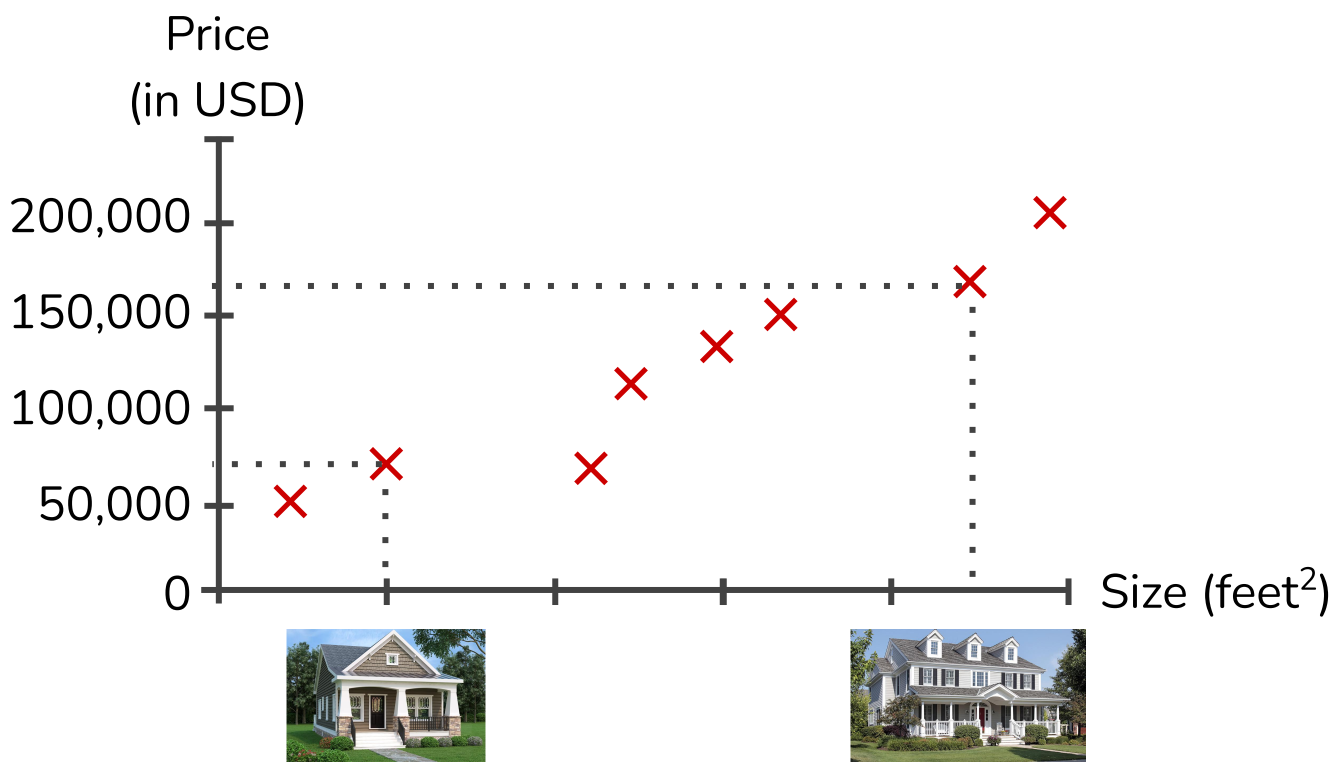
- Estimate the price of a house based on their size.



Linear Regression Example

Problem: House price prediction

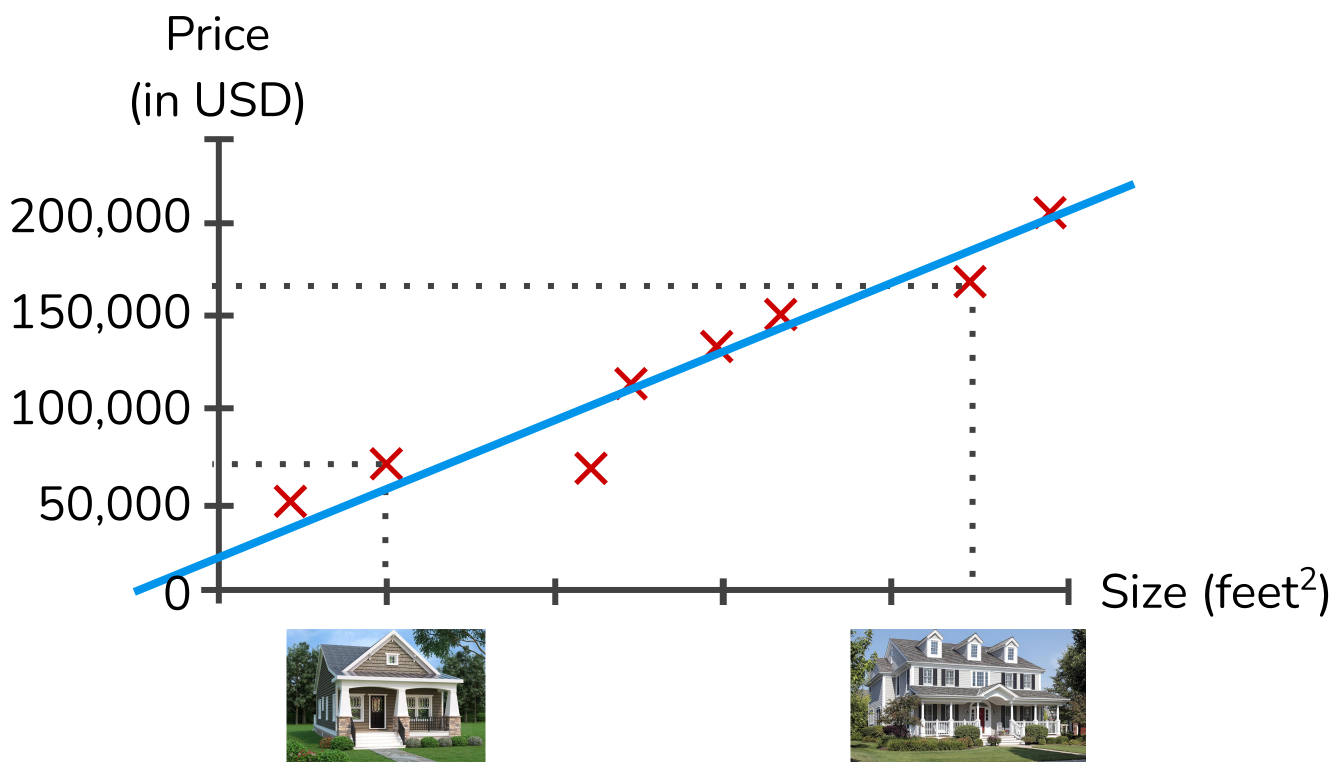
- Estimate the price of a house based on their size.



Linear Regression Example

Problem: House price prediction

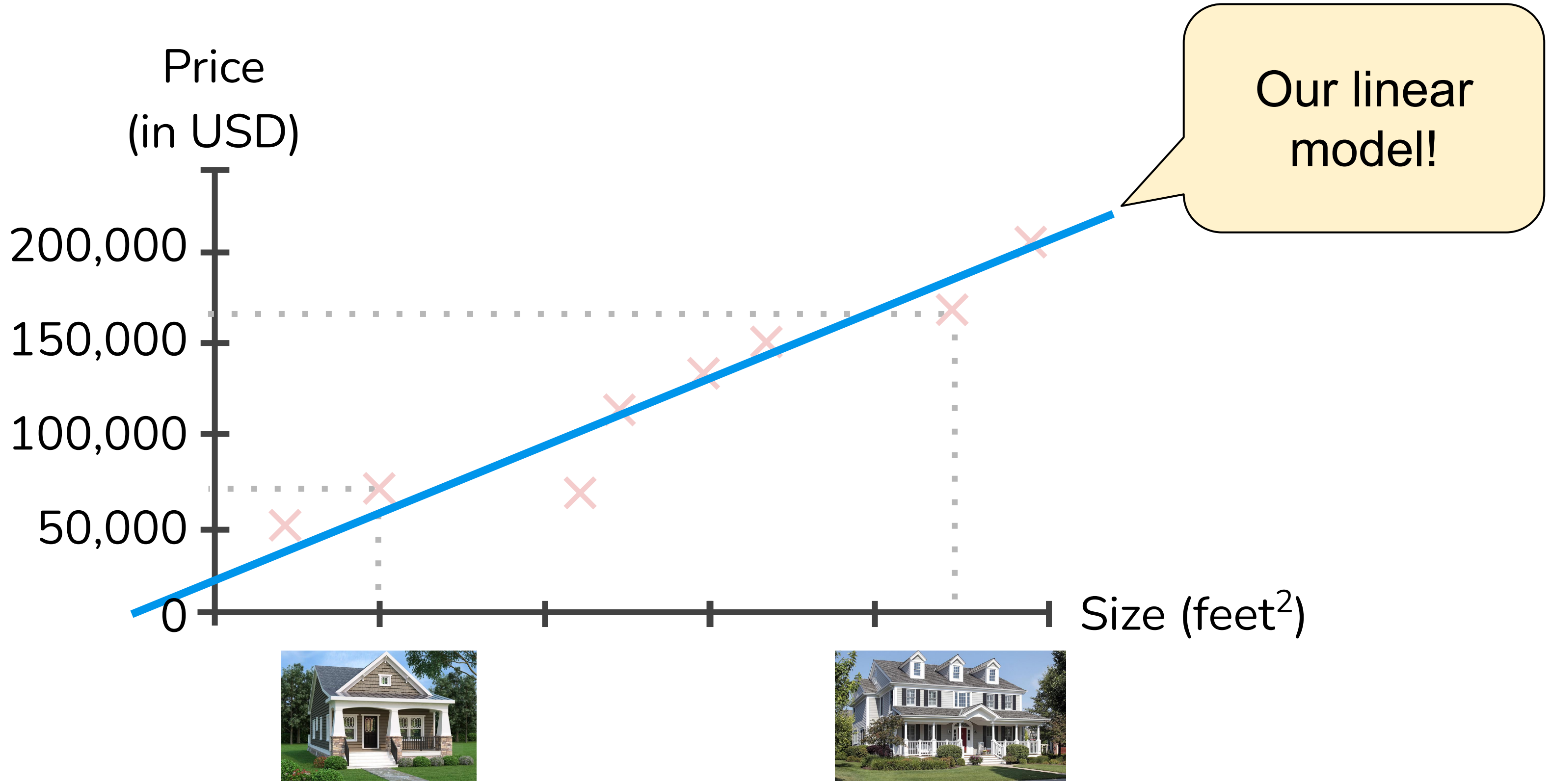
- Estimate the price of a house based on their size.



Linear Regression Example

Problem: House price prediction

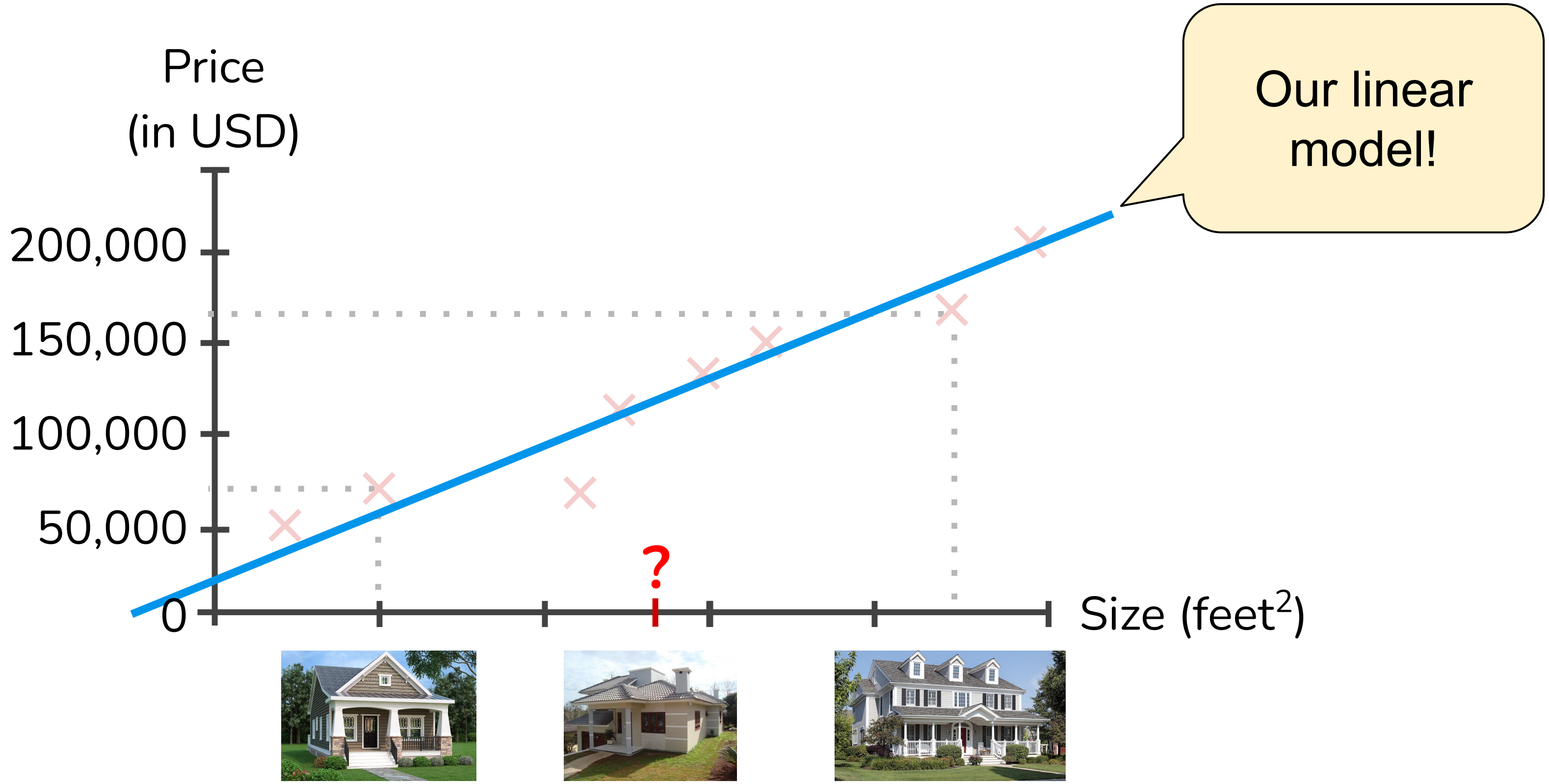
- Estimate the price of a house based on their size.



Linear Regression Example

Problem: House price prediction

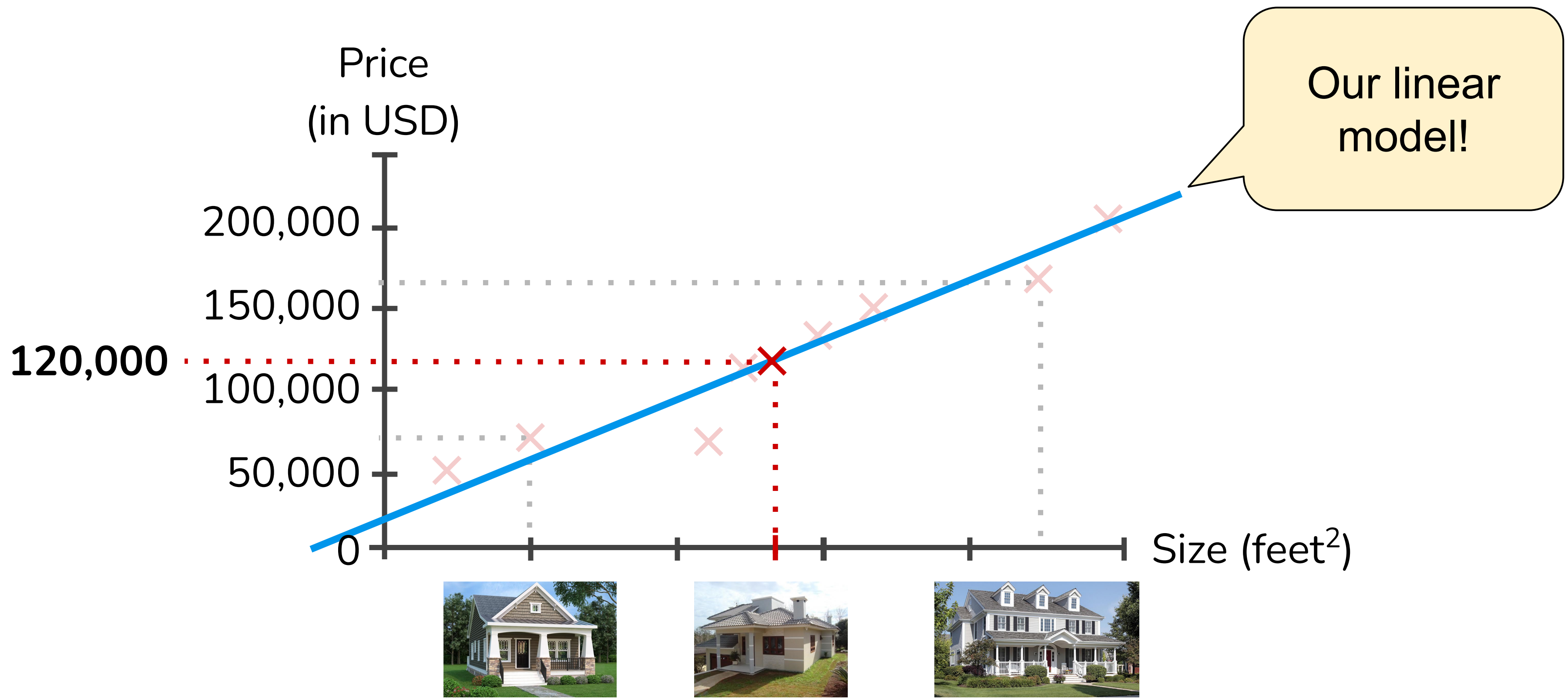
- Estimate the price of a house based on their size.



Linear Regression Example

Problem: House price prediction

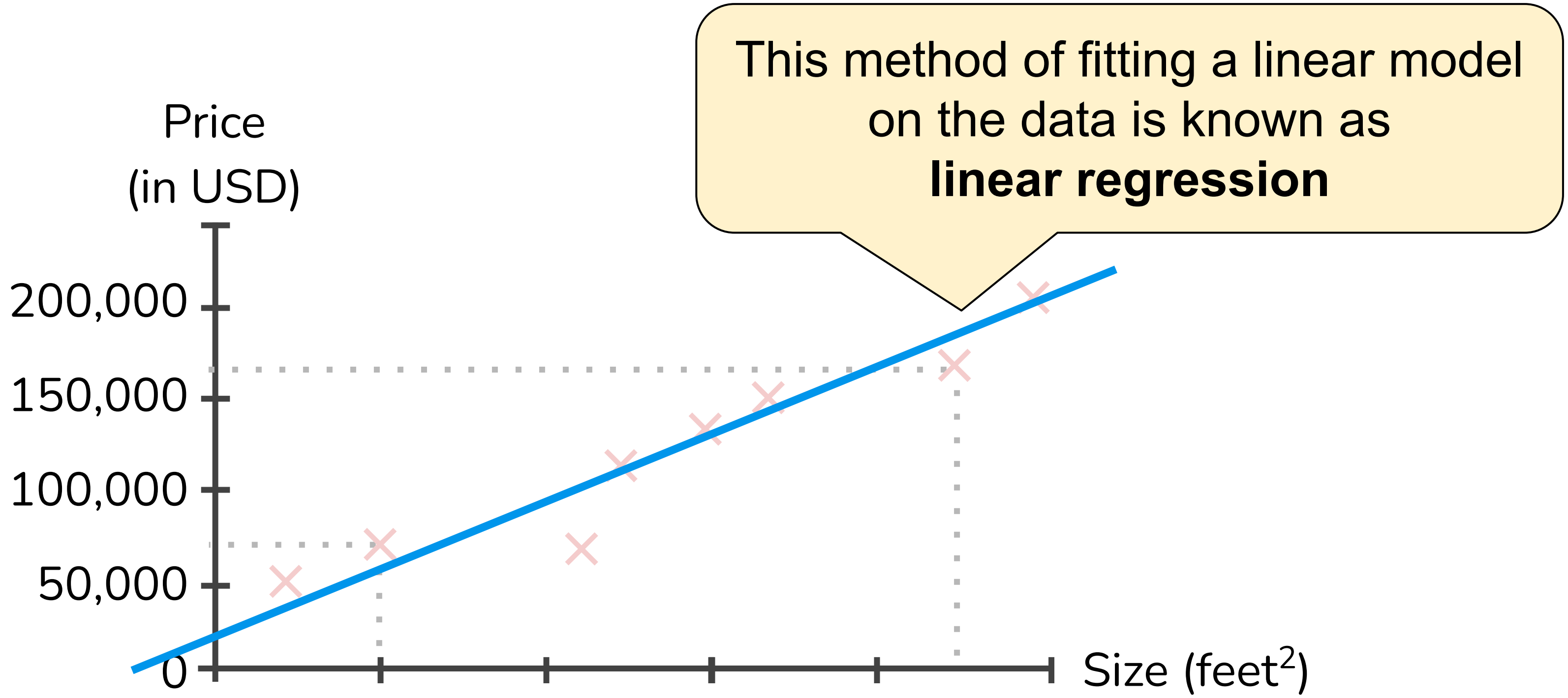
- Estimate the price of a house based on their size.



Linear Regression Example

Problem: House price prediction

- Estimate the price of a house based on their size.





Linear Regression Example

Problem: House price prediction

- Estimate the price of a house based on their size.

Supervised learning: training dataset items (houses) are labeled (houses' prices)

Regression problem: Predict real-valued output (house price)

Model-based learning: Model trained based on the training data

Linear model

Linear model

$$\underline{Price} = \theta_0 + \theta_1 \cdot \underline{Size}$$

Linear model

$$\underline{Price} = \theta_0 + \theta_1 \cdot \underline{Size}$$

Model parameters

Linear Regression Example

Linear model

$$\underline{Price} = \theta_0 + \theta_1 \cdot \underline{Size}$$

Model parameters

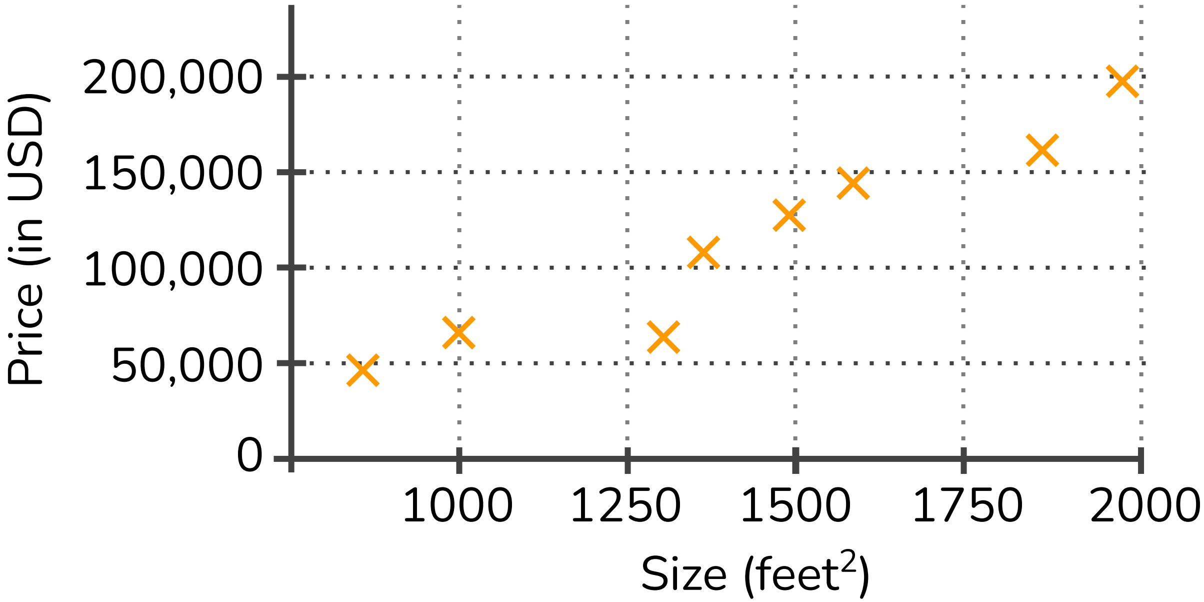
θ_0 is sometimes explained as a bias and represented by the letter b!

Some authors use the term **model weights** instead of model parameters. In this case, they might use w_0, w_1, … to represent the model weights (parameters).

Linear model

$$\underline{Price} = \theta_0 + \theta_1 \cdot \underline{Size}$$

Model parameters



Linear model

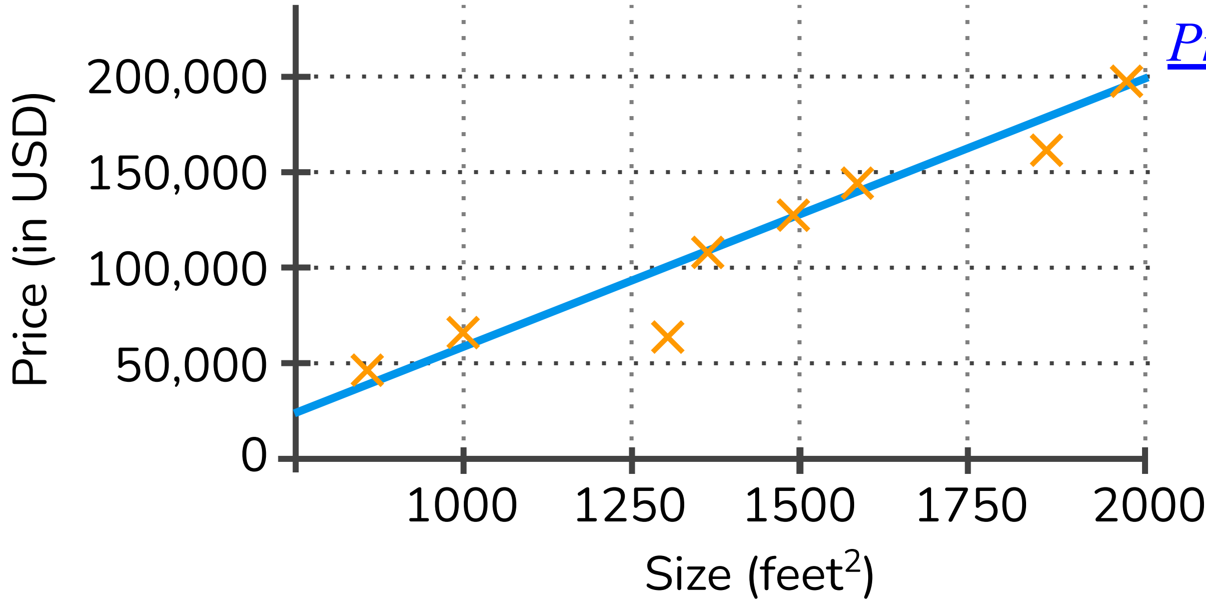
$$\underline{Price} = \theta_0 + \theta_1 \cdot \underline{Size}$$

Model parameters

$\theta_0 = 25000$

$\theta_1 = 87.5$

$\underline{Price} = 25000 + 87.5 \cdot \underline{Size}$



Linear Regression Example

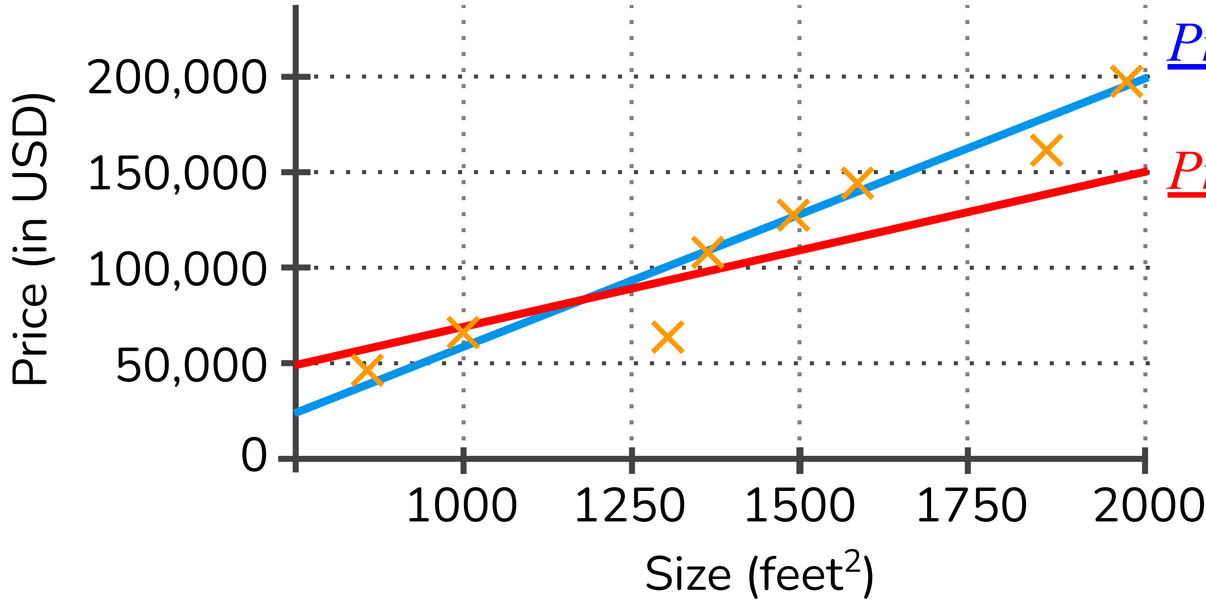
Linear model

$$\underline{Price} = \theta_0 + \theta_1 \cdot \underline{Size}$$

Model parameters

$\theta_0 = 25000$

$\theta_1 = 87.5$



$\underline{Price} = 25000 + 87.5 \cdot \underline{Size}$

$\underline{Price} = 50000 + 50.0 \cdot \underline{Size}$

$\theta_0 = 50000$

$\theta_1 = 50.0$

Linear model terminology

- θ: Set of model parameters. $\theta = \{\theta_0, \theta_1, \ldots, \theta_n\}$

Linear model terminology

- θ: Set of model parameters. $\theta = \{\theta_0, \theta_1, \ldots, \theta_n\}$
- x: Set of input variables. $x = \{x_1, x_2, \ldots, x_n\}$

Only one in our previous example
x_1 = Size

Linear model terminology

- θ: Set of model parameters. $\theta = \{\theta_0, \theta_1, \ldots, \theta_n\}$
- x: Set of input variables. $x = \{x_1, x_2, \ldots, x_n\}$
- $\hat{y}$: model output (predicted value)

In our previous example $\hat{y}$ = Price

Linear model terminology

- θ: Set of model parameters. $\theta = \{\theta_0, \theta_1, \ldots, \theta_n\}$
- x: Set of input variables. $x = \{x_1, x_2, \ldots, x_n\}$
- $\hat{y}$: model output (predicted value)
- $\hat{y} = \theta_0 + \theta_1 \cdot x_1 + \ldots + \theta_n \cdot x_n$ Multivariate linear model
- $\hat{y} = \theta_0 + \theta_1 \cdot x_1$ Univariate linear model

Linear Regression Example

Linear model terminology

- θ: Set of model parameters. $\theta = \{\theta_0, \theta_1, \ldots, \theta_n\}$
- x: Set of input variables. $x = \{x_1, x_2, \ldots, x_n\}$
- $\hat{y}$: model output (predicted value)
- $\hat{y} = \theta_0 + \theta_1 \cdot x_1 + \ldots + \theta_n \cdot x_n$
- $\hat{y} = \theta_0 + \theta_1 \cdot x_1$
- $h_\theta(x) = \theta_0 + \theta_1 \cdot x_1 + \ldots + \theta_n \cdot x_n$

Also known has hypothesis!

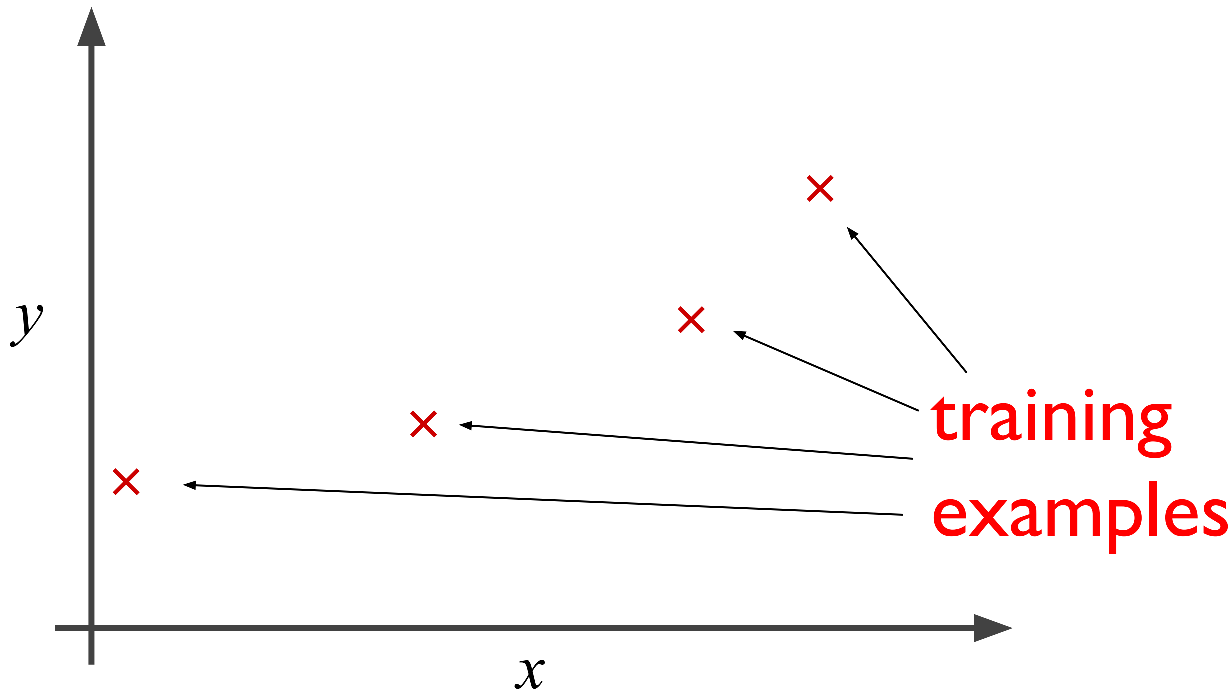
Model training - intuition

Linear Regression Example

Training: choose $\theta = \{\theta_0, \theta_1\}$ so that $h_\theta(x)$ close to y for our training examples (x,y)

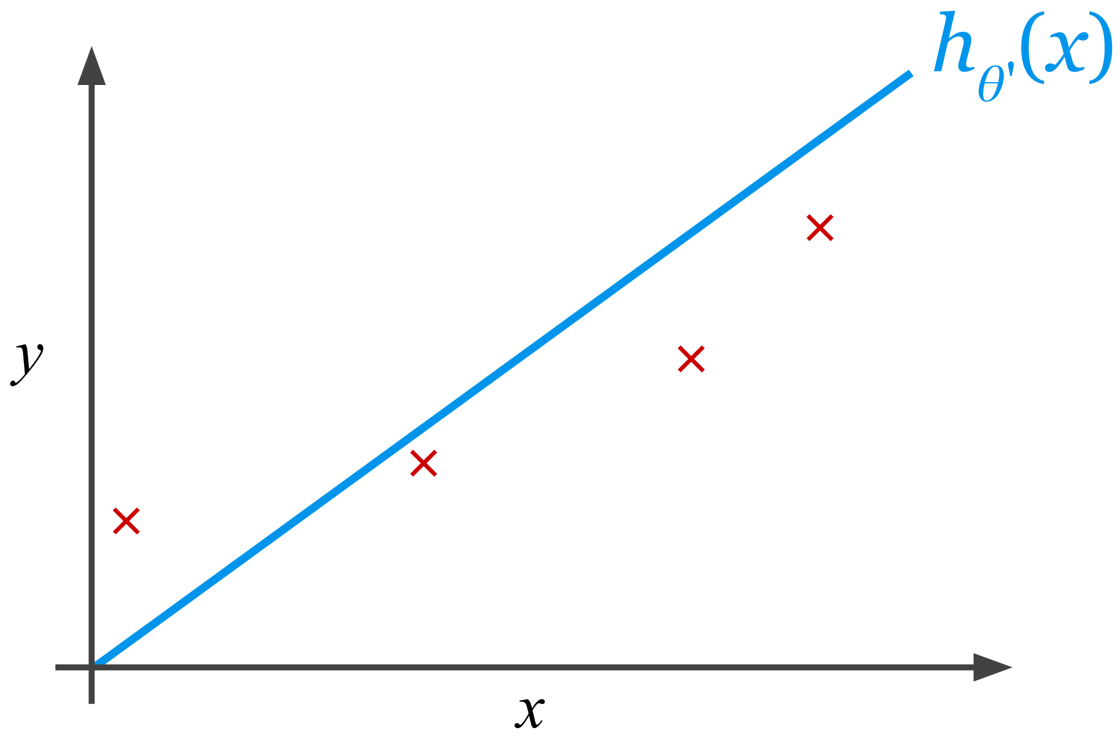
Linear Regression Example

<u>Training</u>: choose $\theta = \{\theta_0, \theta_1\}$ so that $h_\theta(x)$ close to y for our training examples (x,y)



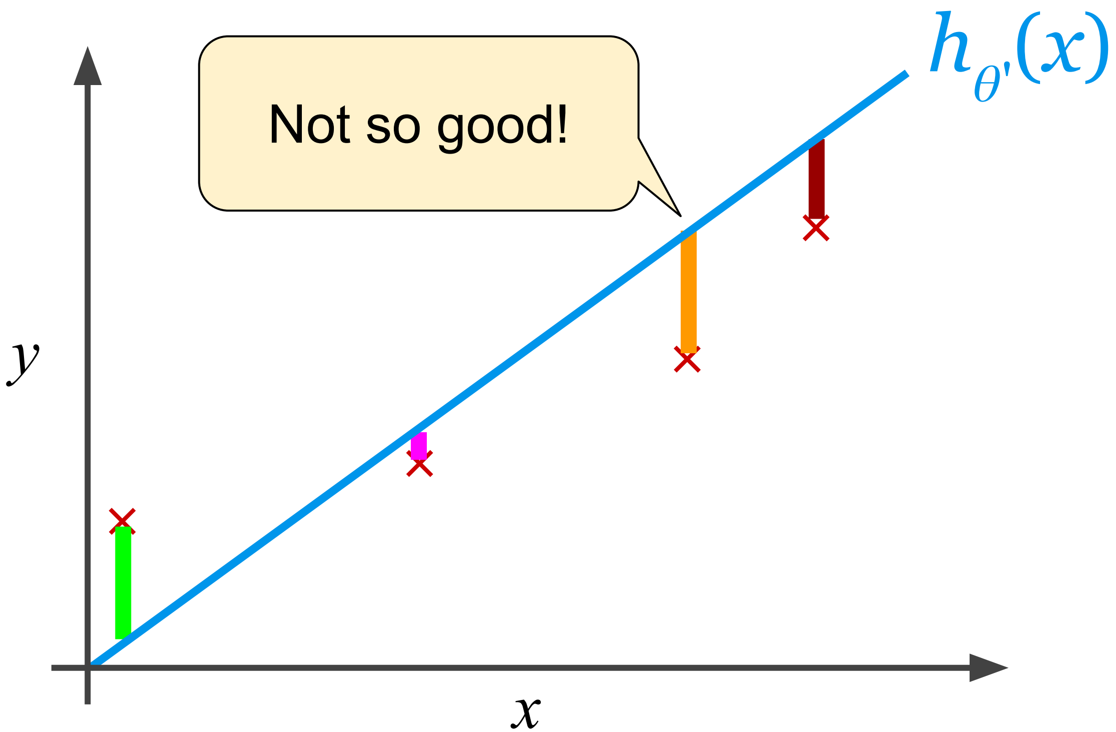
Linear Regression Example

Training: choose $\theta = \{\theta_0, \theta_1\}$ so that $h_\theta(x)$ close to y for our training examples (x,y)



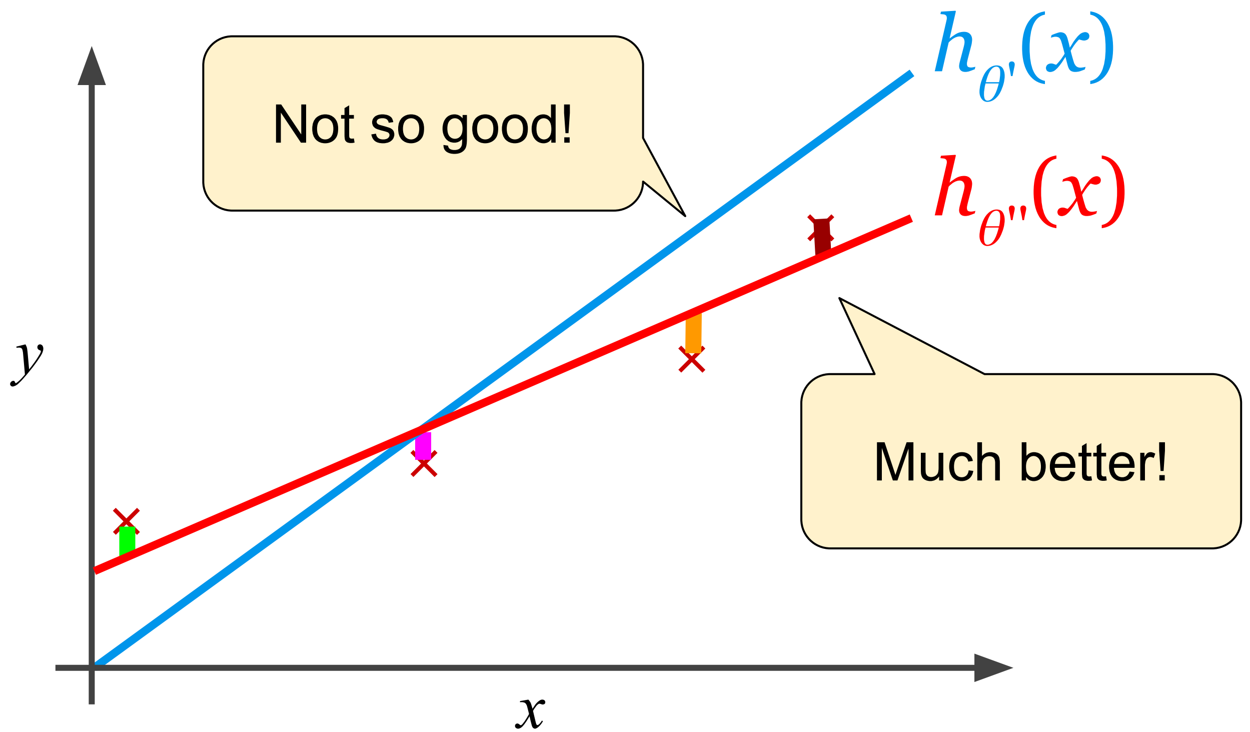
Linear Regression Example

Training: choose $\theta = \{\theta_0, \theta_1\}$ so that $h_\theta(x)$ close to y for our training examples (x,y)



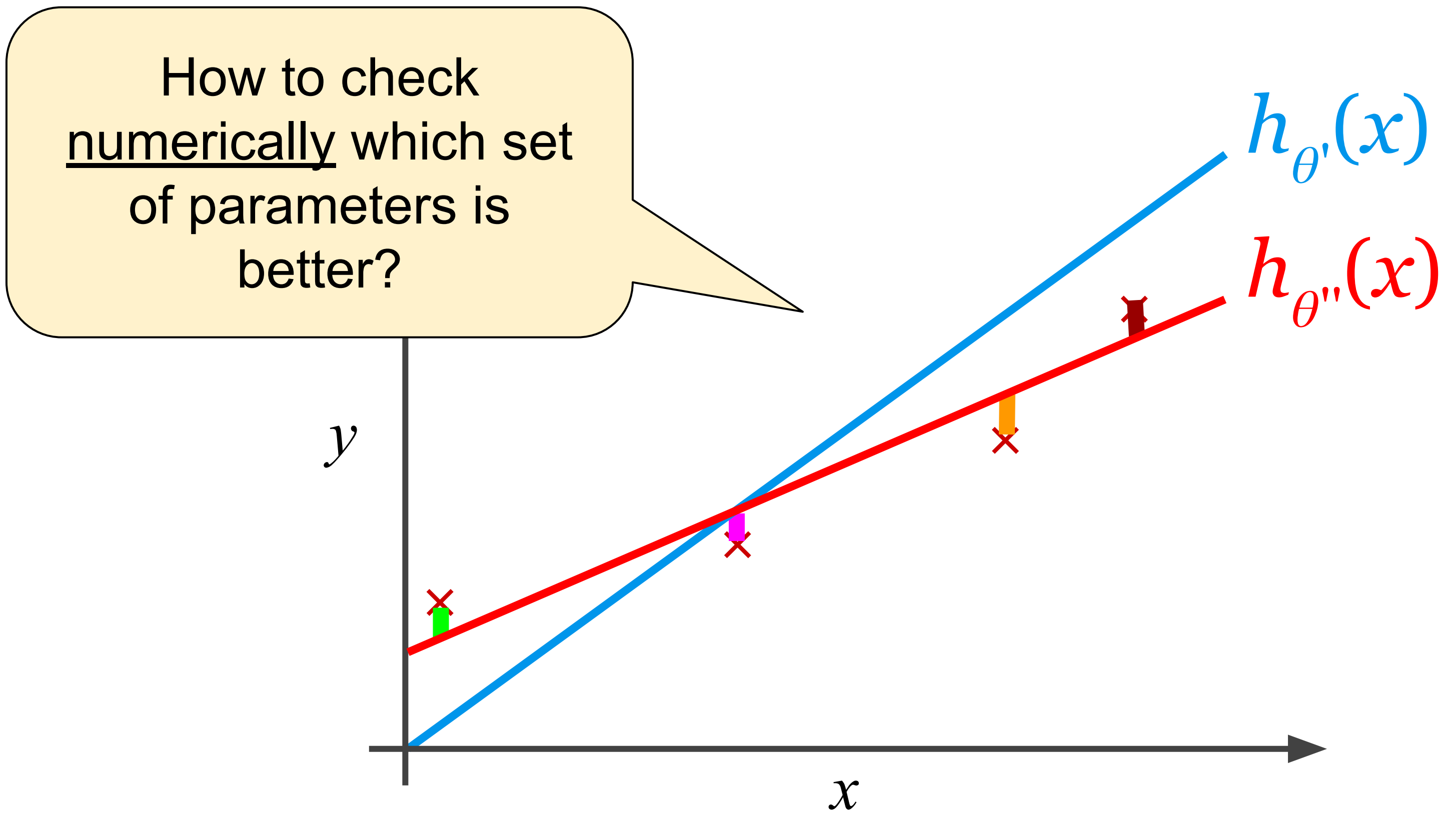
Linear Regression Example

<u>Training</u>: choose $\theta = \{\theta_0, \theta_1\}$ so that $h_\theta(x)$ close to y for our training examples (x,y)



Linear Regression Example

Training: choose $\theta = \{\theta_0, \theta_1\}$ so that $h_\theta(x)$ close to y for our training examples (x,y)

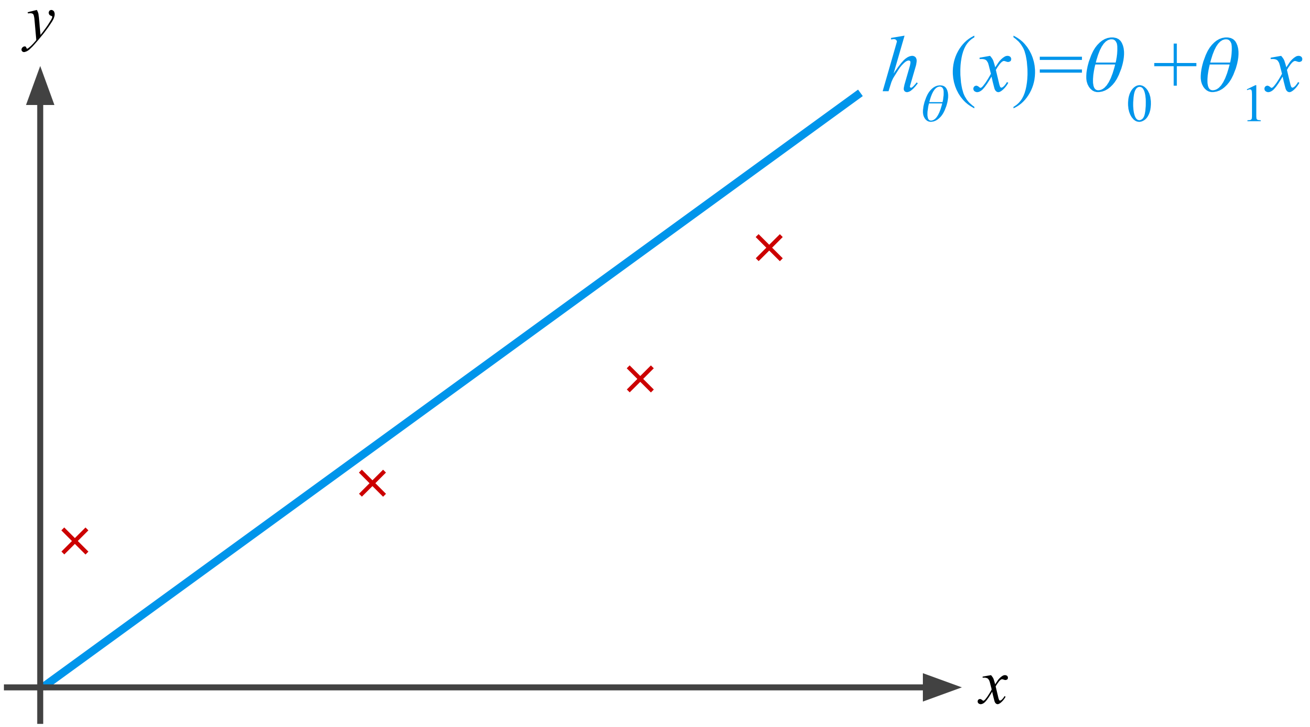


Cost function

Linear Regression Example

Cost function: function that measures how bad the model is

Ex: **Mean squared error**

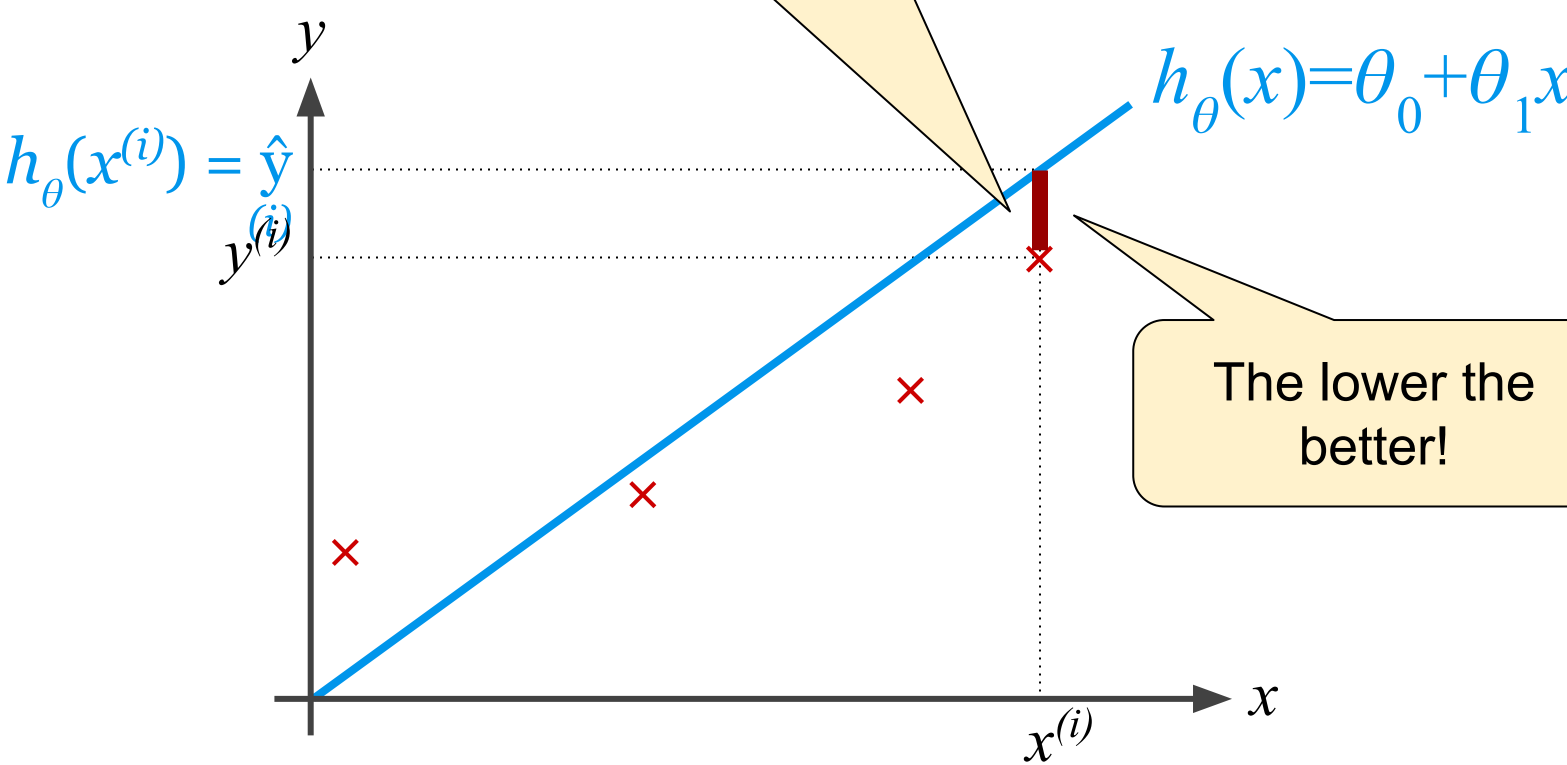


Linear Regression Example

<u>Cost function</u>: fu… res how bad the model is

Ex: **Mean squa…**

Idea: Compute the distance between $\hat{y}^{(i)}$ and $y^{(i)}$, *i.e.*, $h_\theta(x^{(i)}) - y^{(i)}$



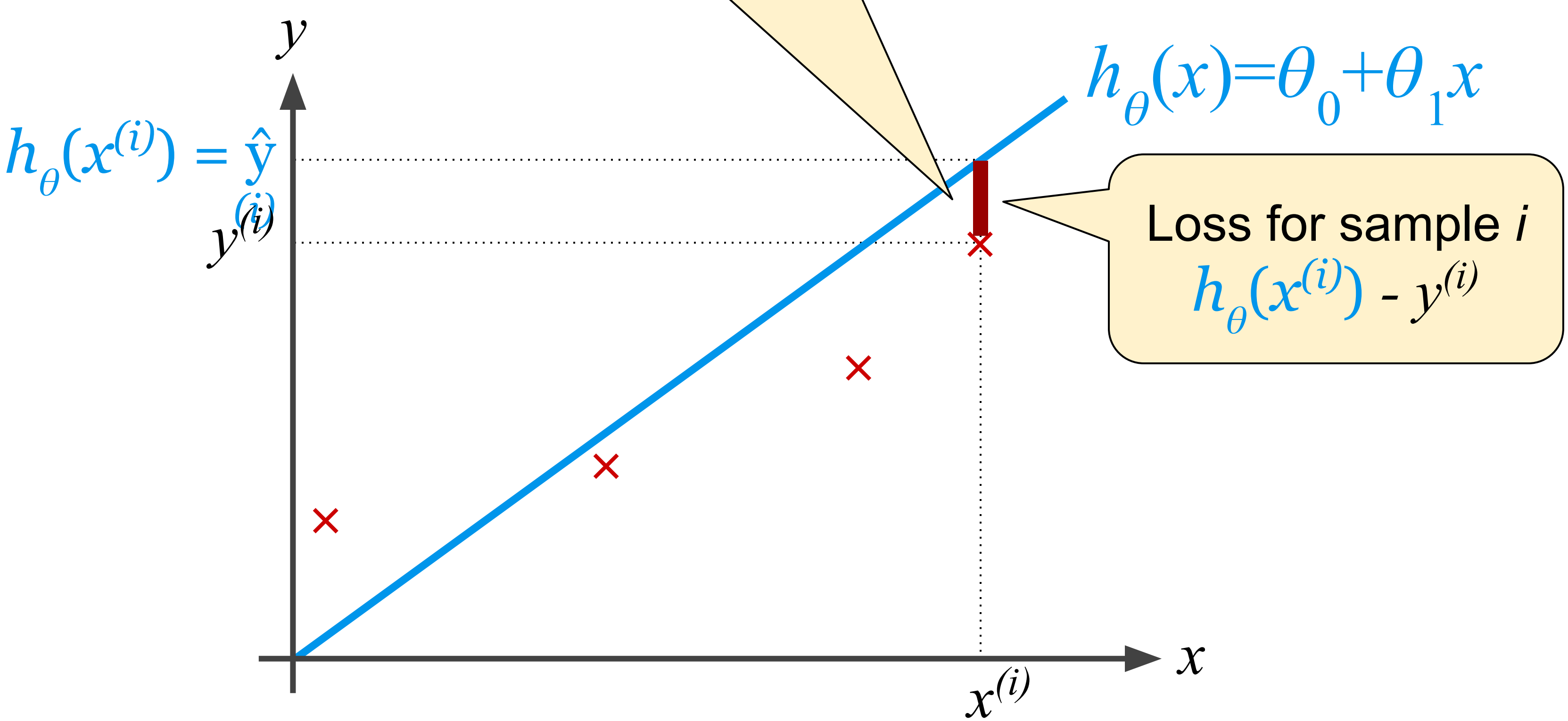
The lower the better!

Linear Regression Example

Cost function: fu… res how bad the model is

Ex: **Mean squa…**

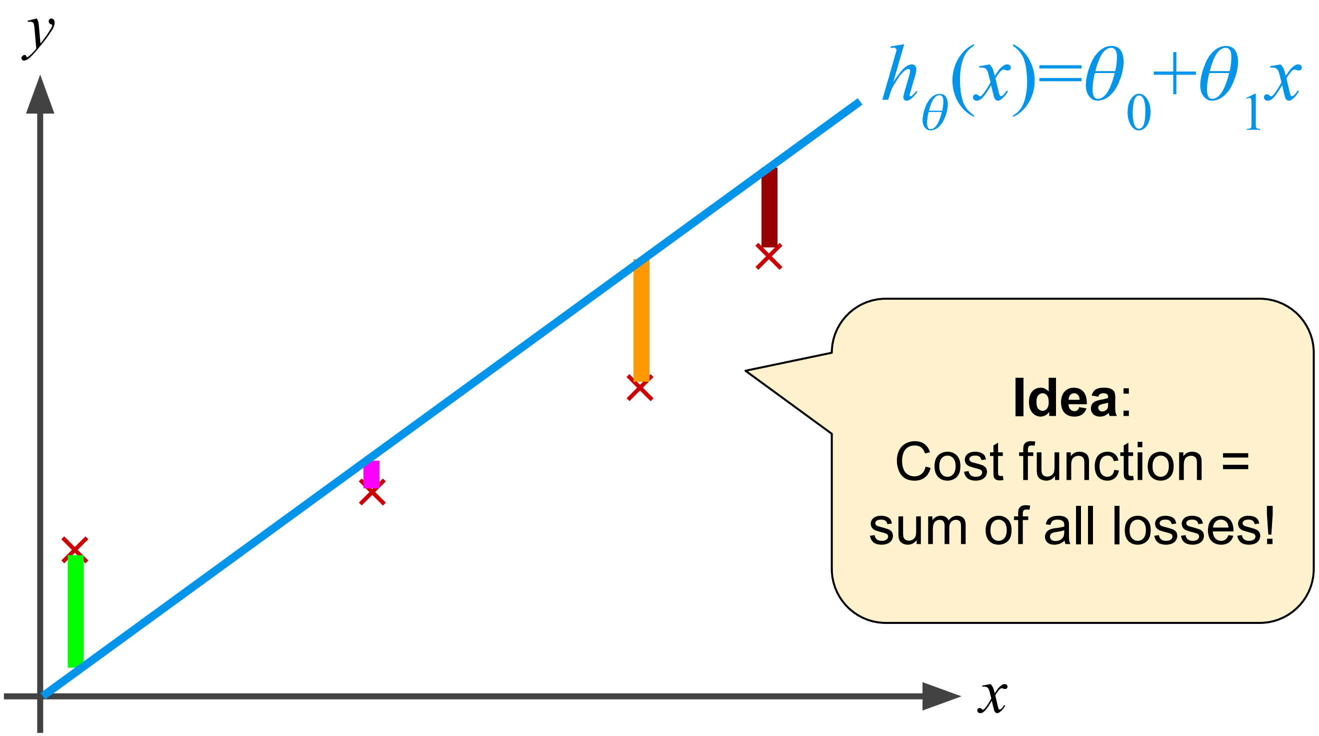
Idea: Compute the distance between $\hat{y}^{(i)}$ and $y^{(i)}$, *i.e.*, $h_\theta(x^{(i)}) - y^{(i)}$



Linear Regression Example

<u>Cost function</u>: function that measures how bad the model is

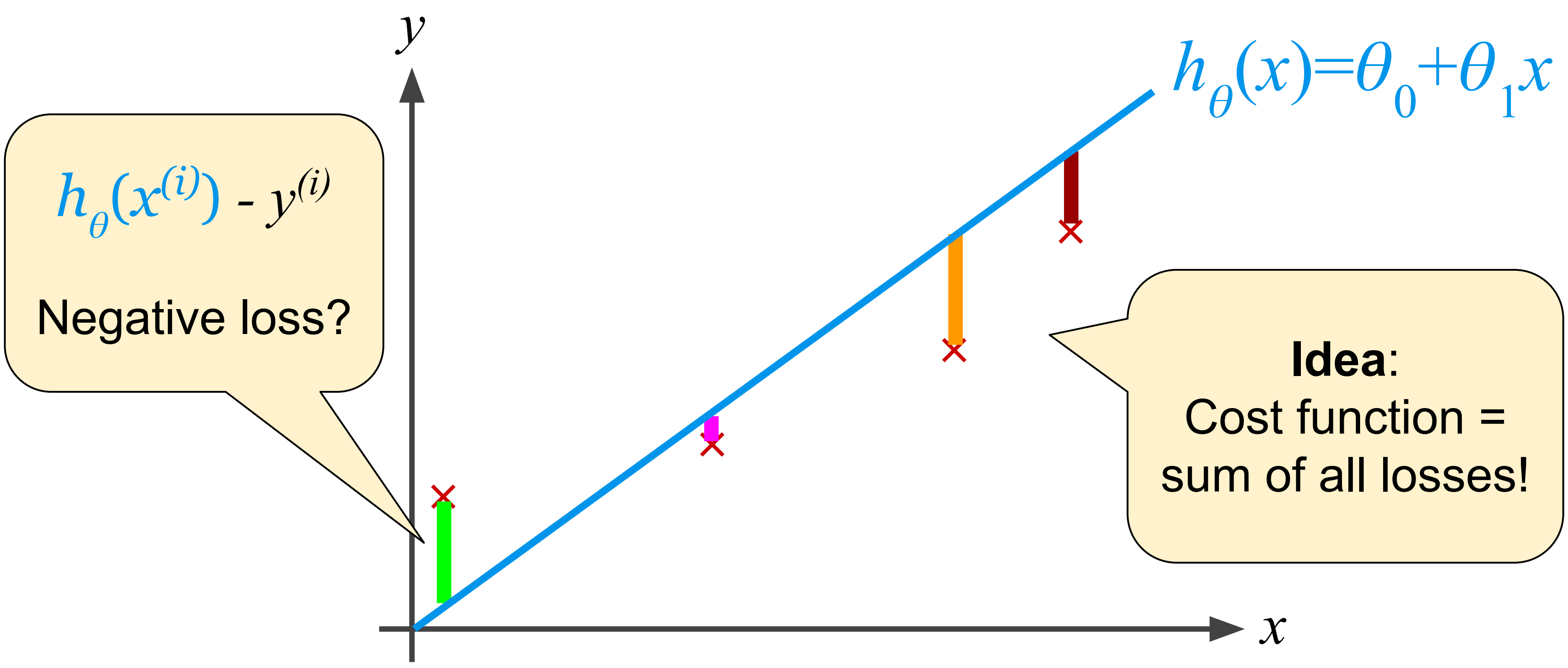
Ex: **Mean squared error**



Linear Regression Example

Cost function: function that measures how bad the model is

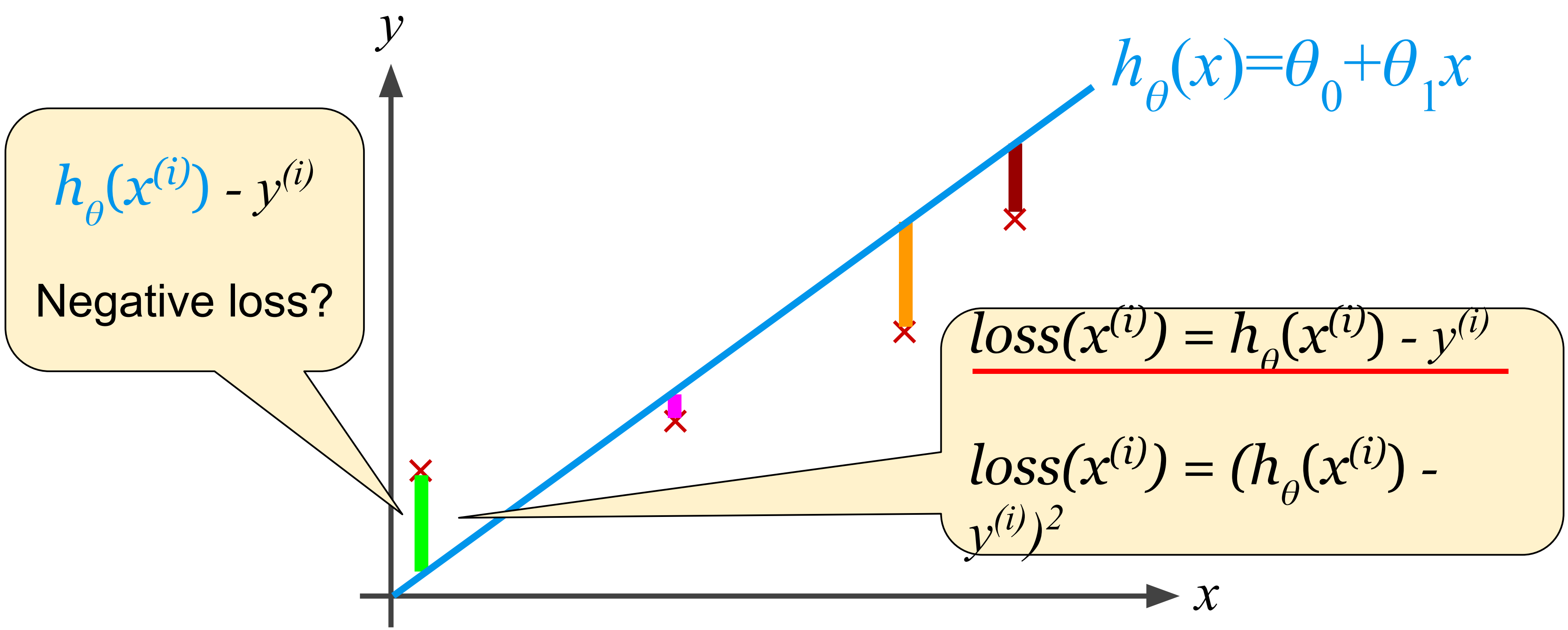
Ex: **Mean squared error**



Linear Regression Example

<u>Cost function</u>: function that measures how bad the model is

Ex: **Mean squared error**



Cost function: function that measures how bad the model is

Ex: **Mean squared error**

Let:

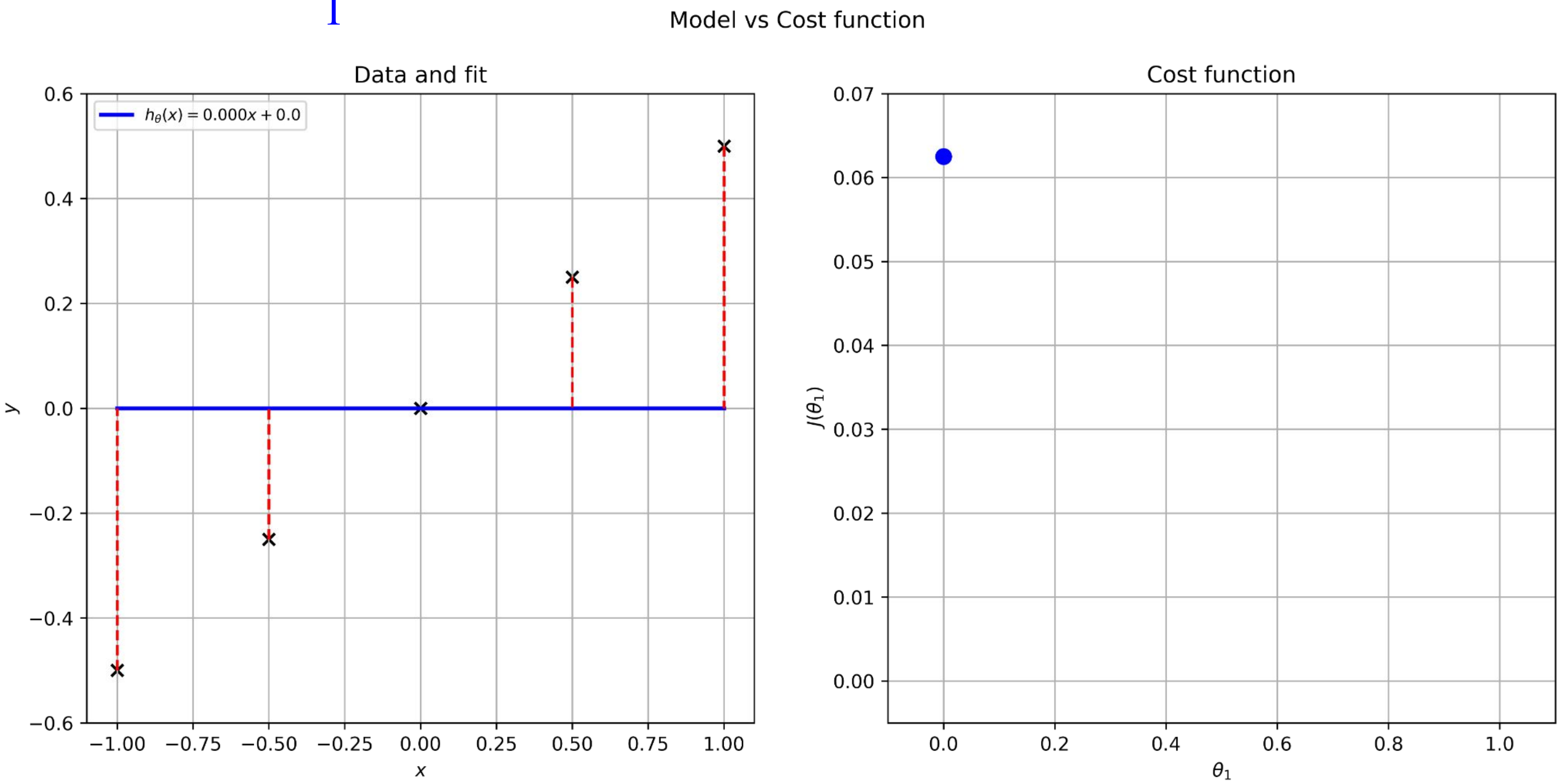
- $X = \{x^1, x^2, \ldots, x^m\}$: be features of the training examples
- $Y = \{y^1, y^2, \ldots, y^m\}$: be the training examples labels
- $\theta = \{\theta_0, \theta_1\}$: be the model parameters

$$\mathrm{MSE}(X, h_\theta) = \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2$$

Linear Regression Example

Effect of θ_1 on $\mathrm{MSE}(X, h_\theta)$

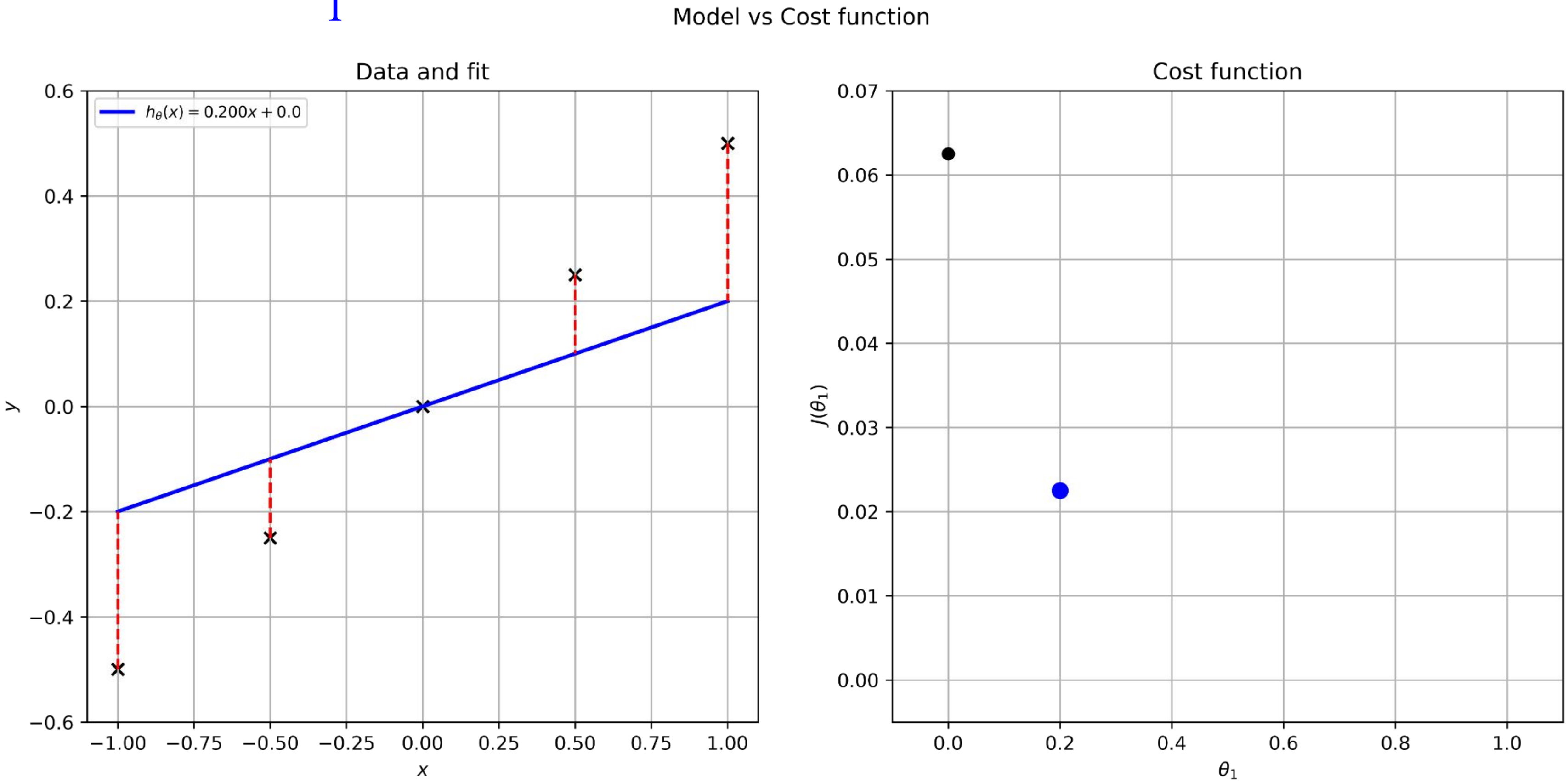
$\theta_1 = 0.0$



Linear Regression Example

Effect of θ_1 on $\mathrm{MSE}(X, h_\theta)$

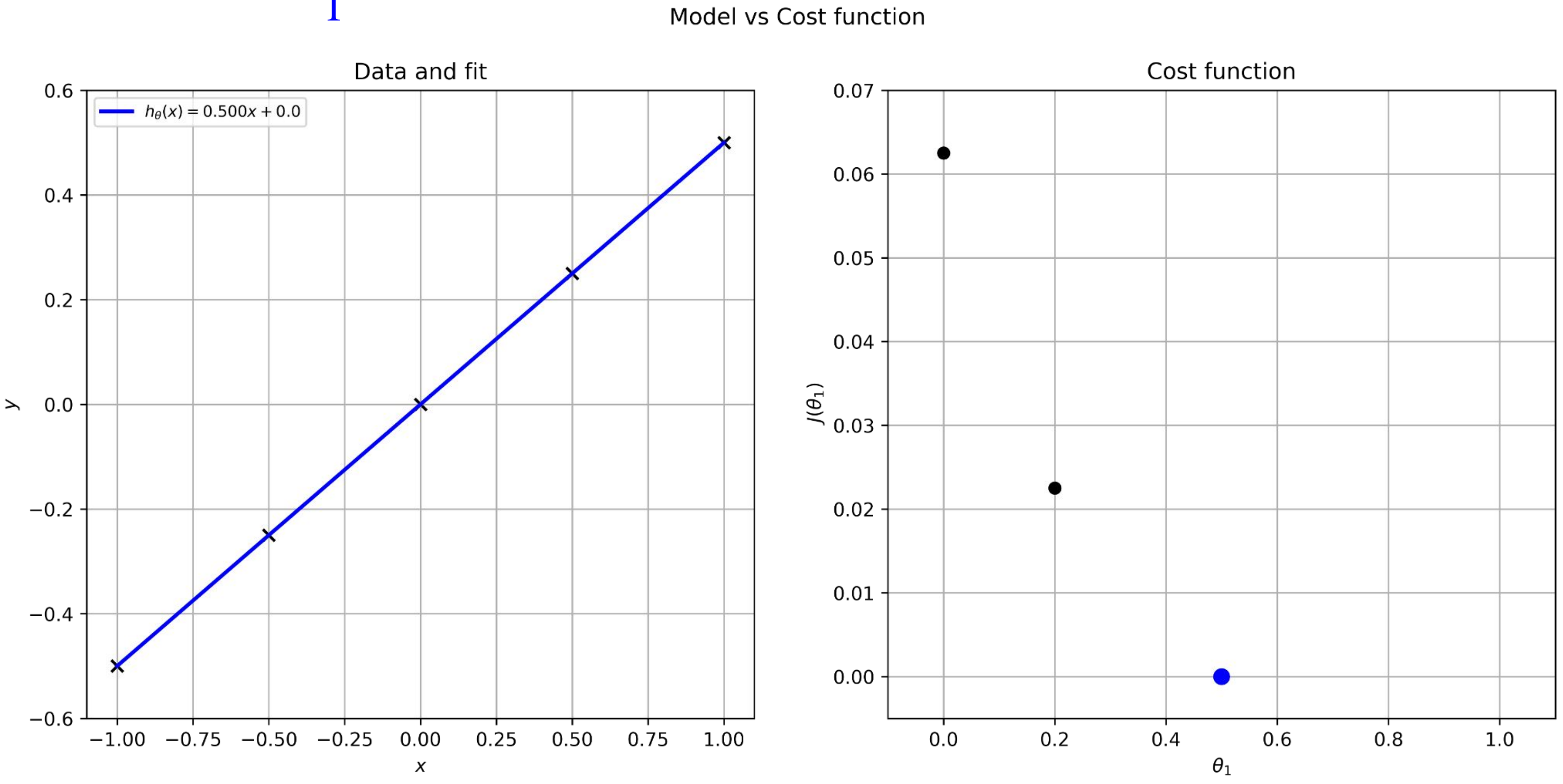
$\theta_1 = 0.2$



Linear Regression Example

Effect of θ_1 on $\mathrm{MSE}(X, h_\theta)$

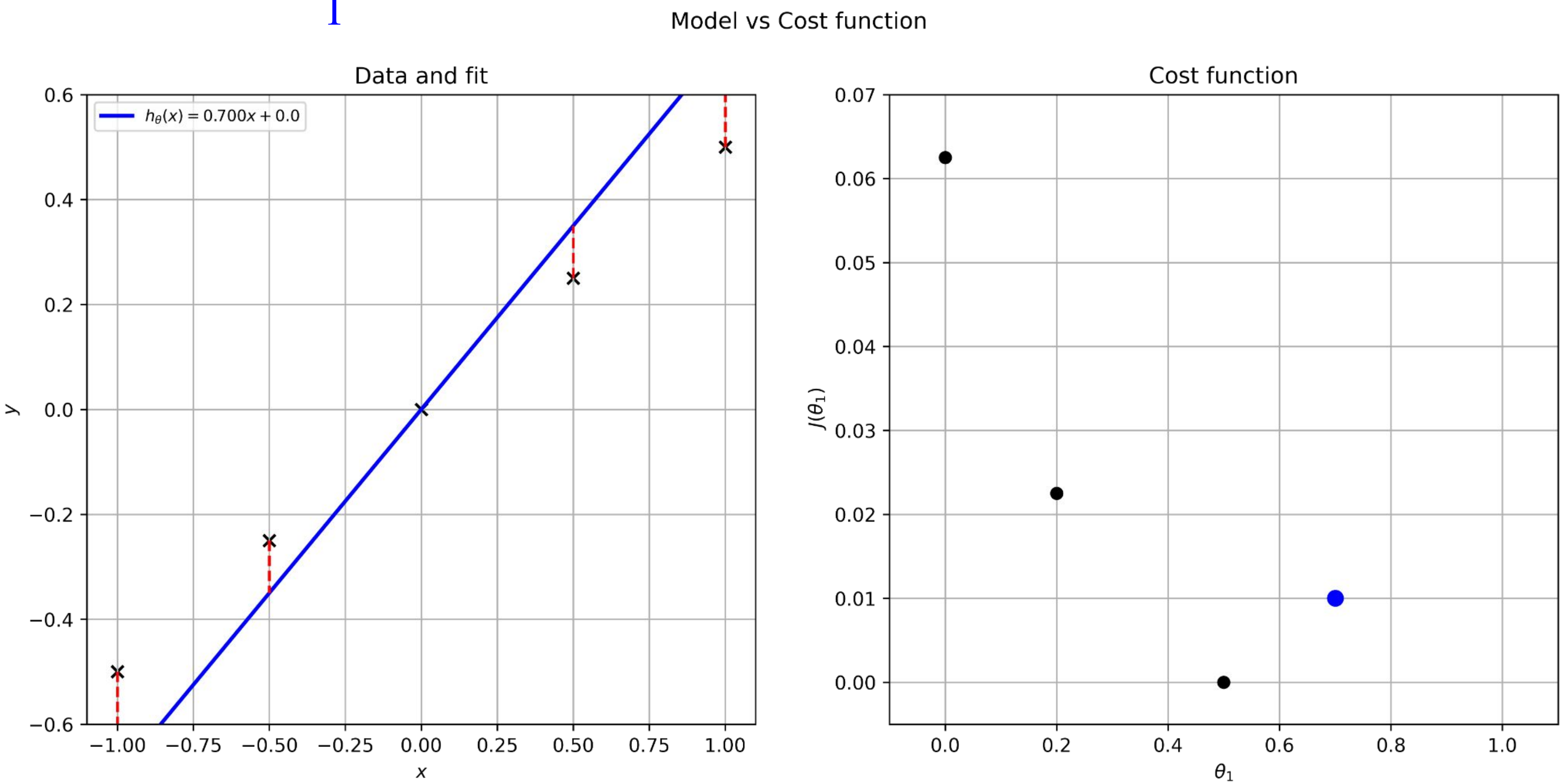
$\theta_1 = 0.5$



Linear Regression Example

Effect of θ_1 on $\mathrm{MSE}(X, h_\theta)$

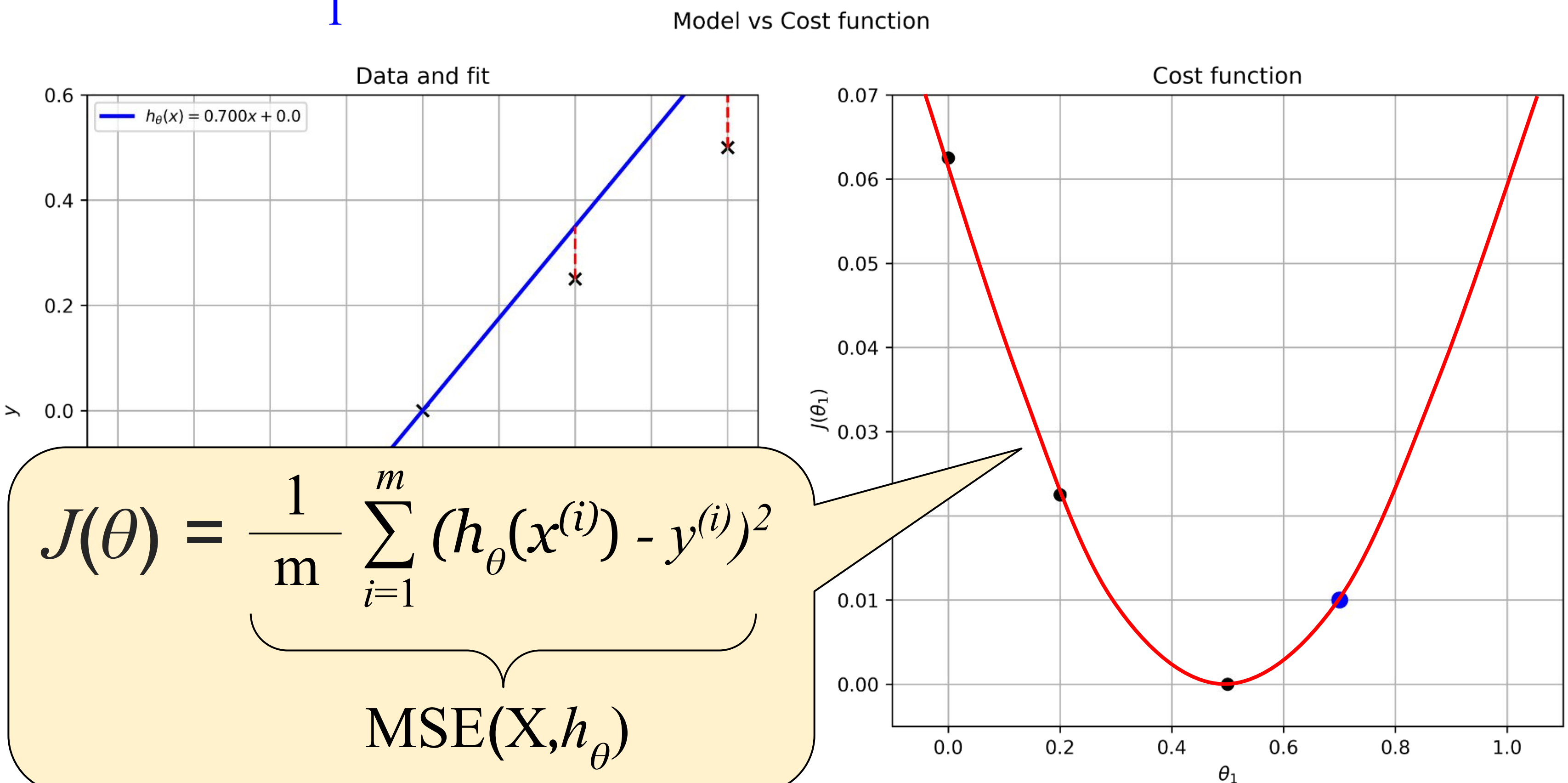
$\theta_1 = 0.7$



Linear Regression Example

Effect of θ_1 on $\mathrm{MSE}(X, h_\theta)$

$\theta_1 = 0.7$



$$J(\theta) = \underbrace{\frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2}_{\mathrm{MSE}(X, h_\theta)}$$

Model training - formalization

Linear Regression Example

<u>Training the model</u>: choose values for $\theta = \{\theta_0, \theta_1\}$ so that the cost function is minimized!

Ex: Train the model using the MSE as cost function

$$\underset{\theta_0,\theta_1}{\text{minimize}} \underbrace{\frac{1}{\mathrm{m}} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})^2}_{\mathrm{MSE}(\mathrm{X}, h_\theta)}$$

Linear Regression Example

Training the model: choose values for $\theta = \{\theta_0, \theta_1\}$ so that the cost functio

Ex: Train the model u

There are several methods to solve this problem. **Gradient Descent** is one of them.

$$\underset{\theta_0,\theta_1}{\text{minimize}} \underbrace{\frac{1}{m}\sum_{i=1}^{m}(h_\theta(x^{(i)}) - y^{(i)})^2}_{\text{MSE}(X,h_\theta)}$$

Instituto de Computação

UNIVERSIDADE ESTADUAL DE CAMPINAS

Capacitação profissional em tecnologias de Inteligência Artificial

Machine Learning Overview

Prof. Edson Borin

https://www.ic.unicamp.br/~edson

Institute of Computing - UNICAMP