A Gentle Introduction to Machine Learning: Basics of Linear Regression

Elie Kawerk

Outline

- What is Machine Learning?
- Types of Machine Learning.
- Linear Regression Objective.
- Cost Function and the Gradient Descent Algorithm.
- Linear Regression and Gradient Descent.

Machine Learning

Machine learning is the subfield of computer science that, according to Arthur Samuel in 1959, gives "computers the ability to learn without being explicitly programmed.“

- Wikipedia

Applications of Machine Learning

- Spam filtering (email)
- Recommender systems (Amazon, Facebook, …)
- Stock market prediction
- Churn prediction
- Computer vision
- Speech recognition
- Fraud detection
- and more ……..

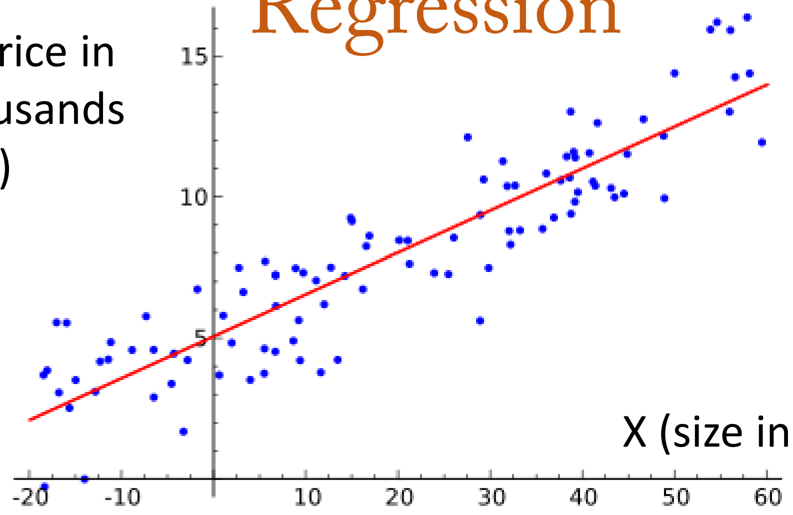
Types of Machine Learning

Supervised learning: features (X) and labels (Y) are given.
The task is to learn the mapping f between X and Y.

$$f\colon X \rightarrow Y$$

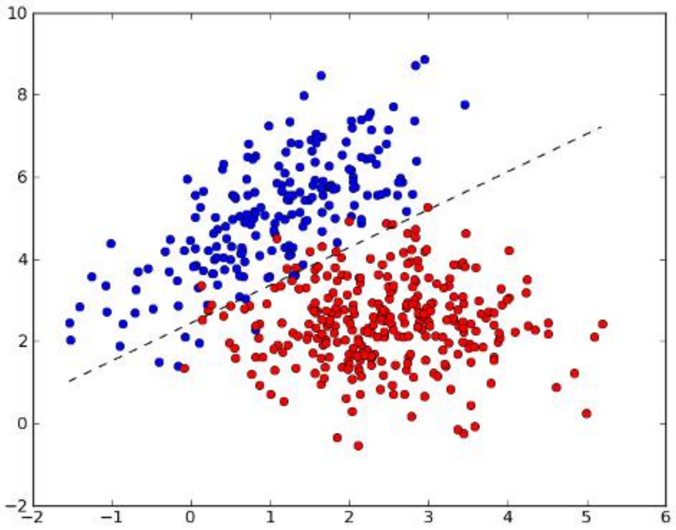
Regression

Y (price in thousands of $)



X (size in m^2)

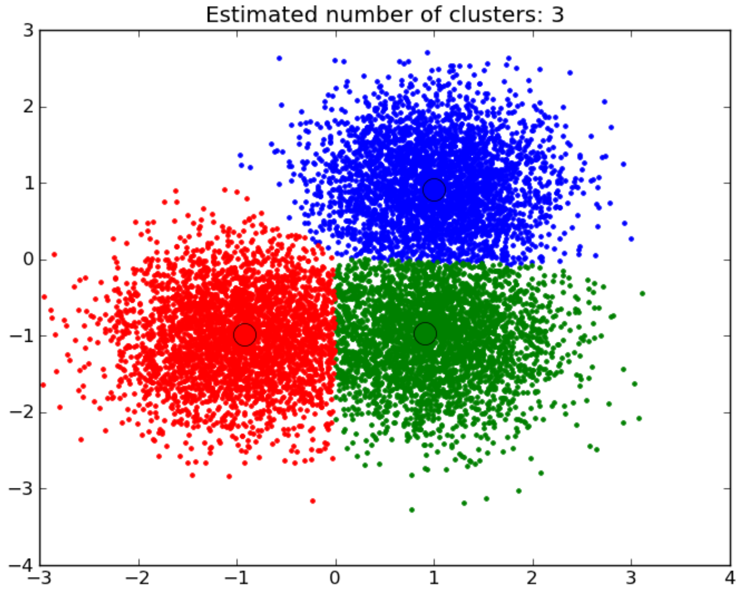
Classification



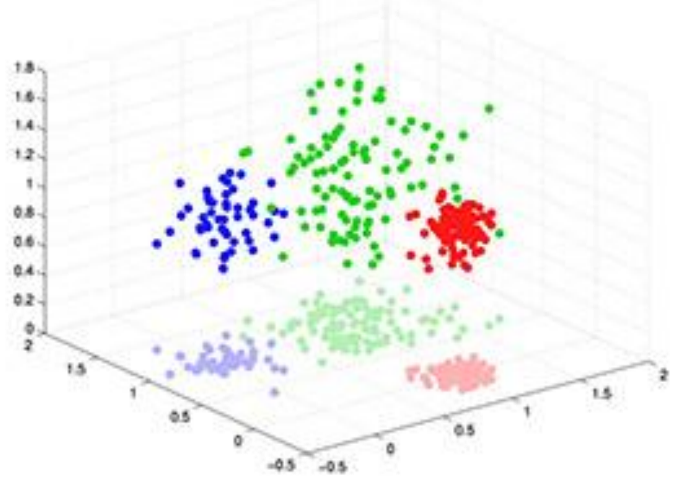
Types of Machine Learning

Unsupervised learning: only features (X) are provided.

Clustering

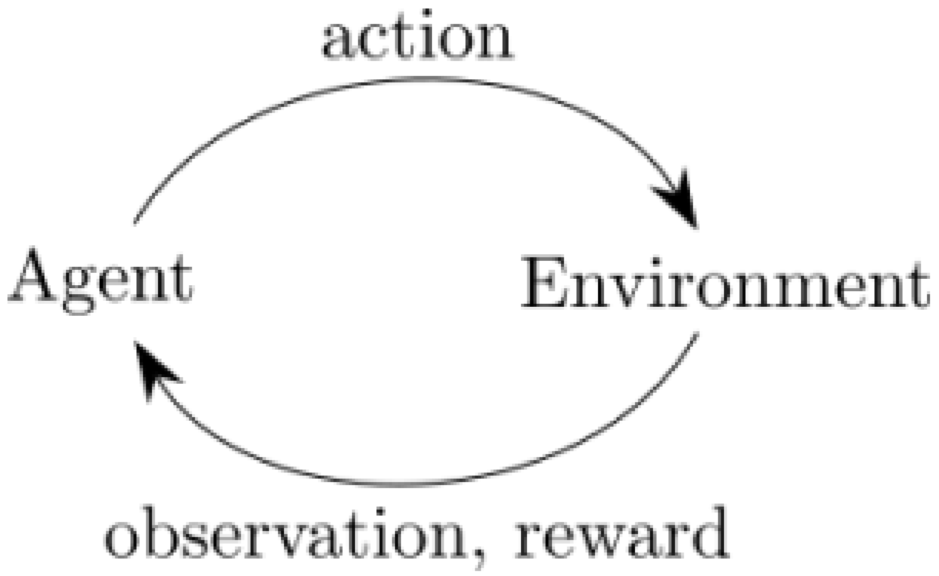


Dimensionality reduction



Types of Machine Learning

Reinforcement learning



Linear Regression

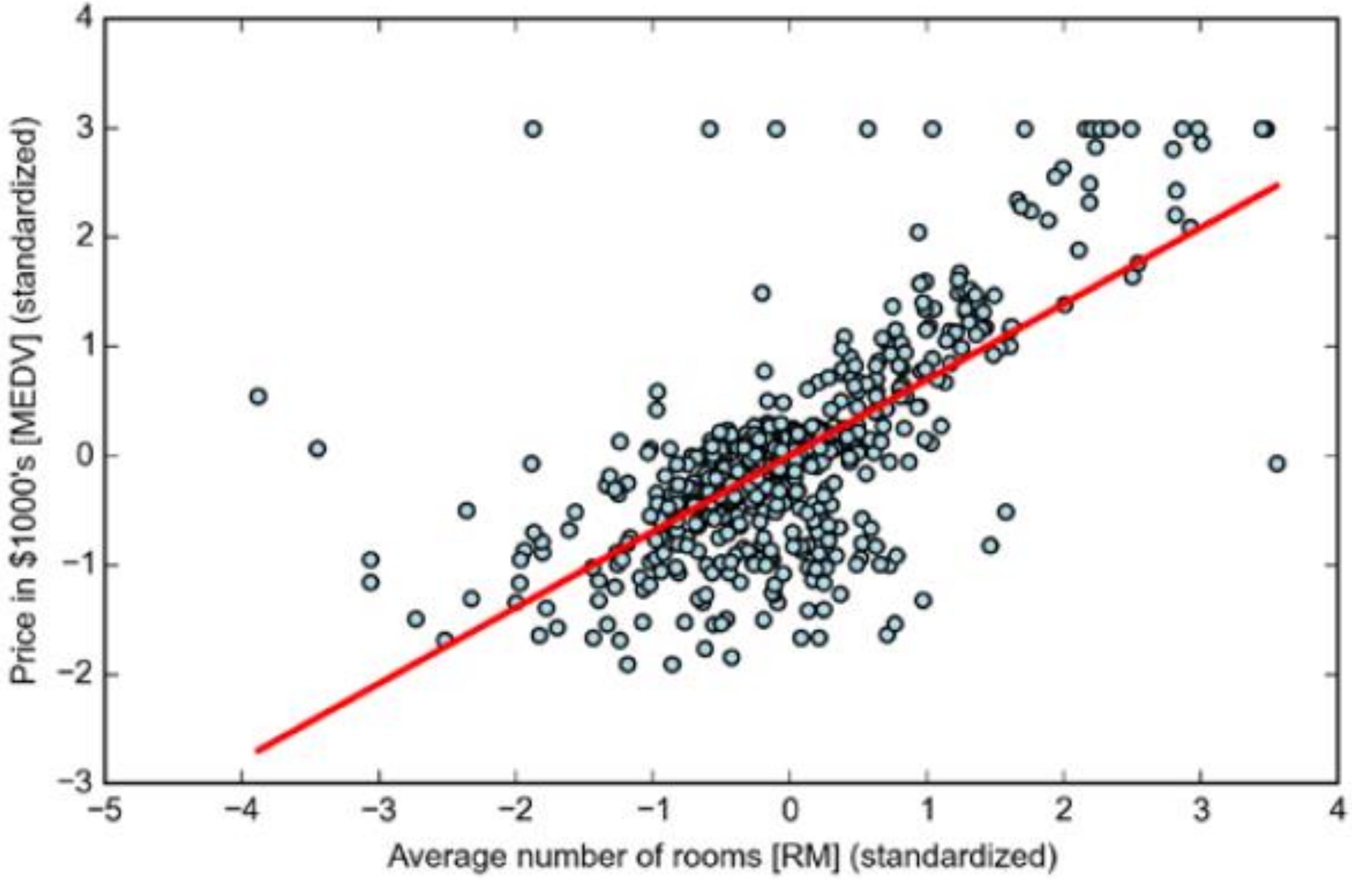
Given the Boston housing dataset (UCI machine learning repository).

X (Average Number of Rooms)

Y (Median House Price)

	CRIM	ZN	INDUS	CHAS	NOX	RM	AGE	DIS	RAD	TAX	PTRATIO	B	LSTAT	MEDV
0	0.00632	18	2.31	0	0.538	6.575	65.2	4.0900	1	296	15.3	396.90	4.98	24.0
1	0.02731	0	7.07	0	0.469	6.421	78.9	4.9671	2	242	17.8	396.90	9.14	21.6
2	0.02729	0	7.07	0	0.469	7.185	61.1	4.9671	2	242	17.8	392.83	4.03	34.7
3	0.03237	0	2.18	0	0.458	6.998	45.8	6.0622	3	222	18.7	394.63	2.94	33.4
4	0.06905	0	2.18	0	0.458	7.147	54.2	6.0622	3	222	18.7	396.90	5.33	36.2

Linear Regression



Linear Regression

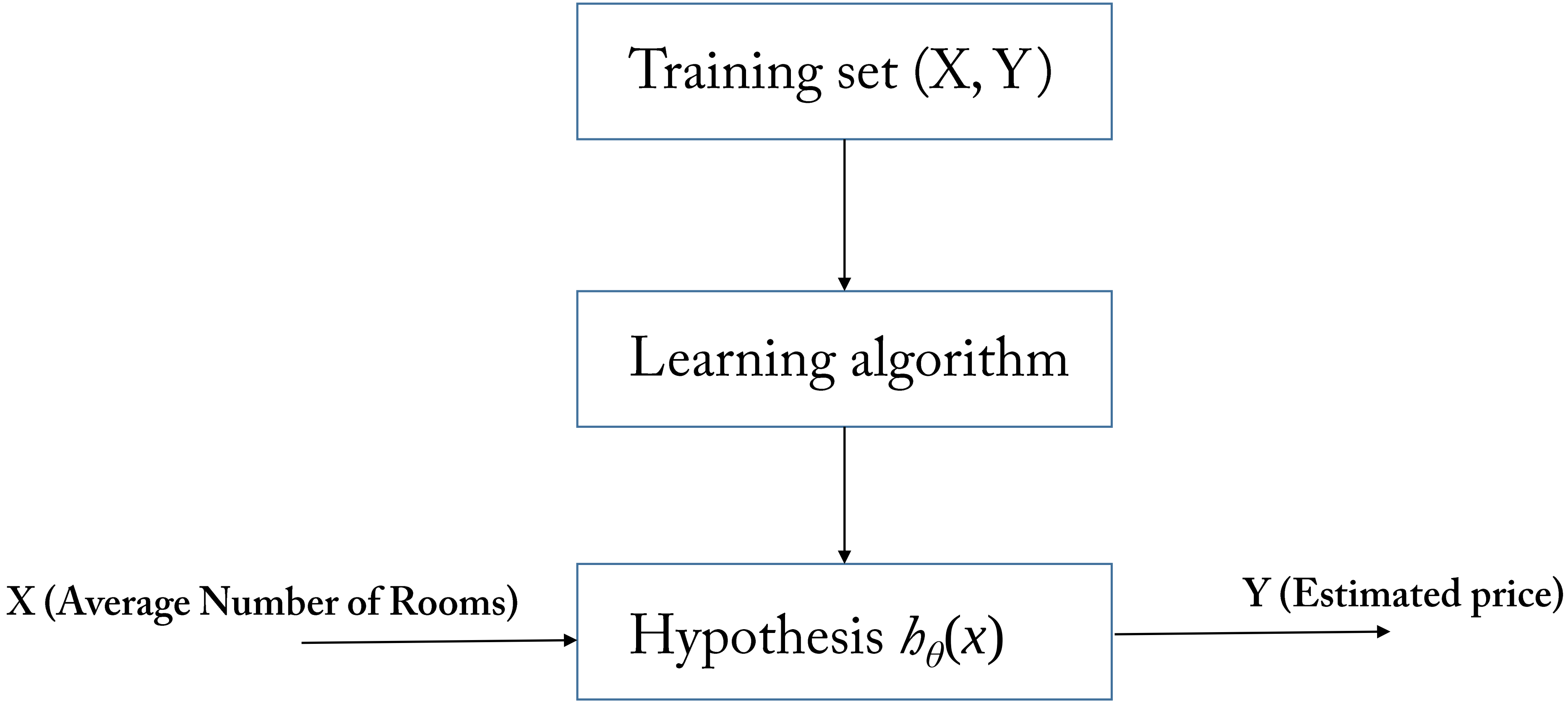
Model

In 1D the linear regression model $h_\theta(x)$ (also known as the hypothesis) has the following form:

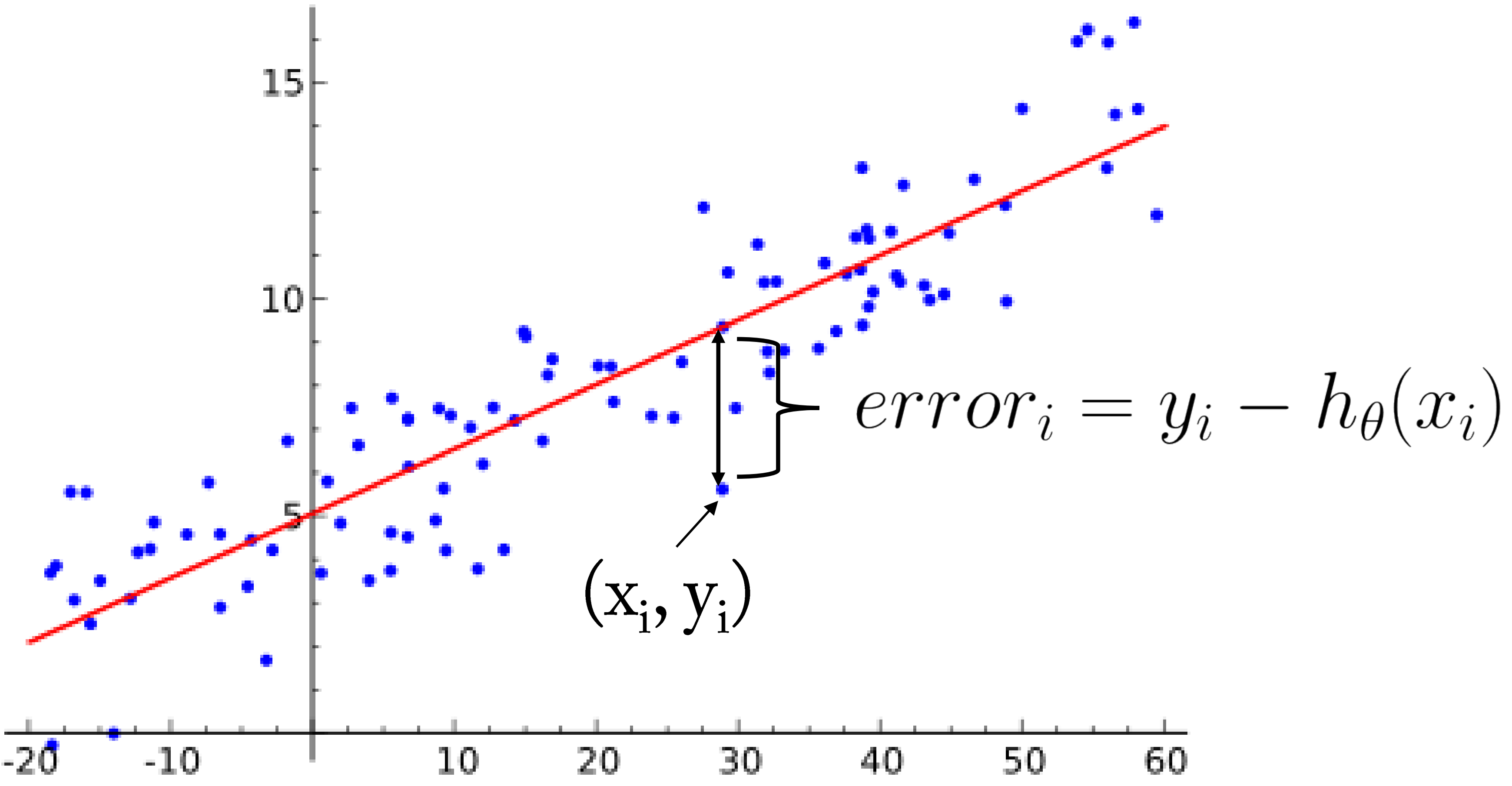
$$h_\theta(x) = \underbrace{\theta_0}_{intercept} + \underbrace{\theta_1}_{slope} \times x$$

The task is to determine the θ_i 's with $\theta = (\theta_0, \theta_1)$

Linear Regression



Cost function



Cost function

A cost function measures the agreement between the true labels and the predicted ones. Given a dataset of N observations.

$$((x_1, y_1), \dots, (x_i, y_i), \dots, (x_N, y_N))$$

The cost function for linear regression is the mean squared error:

$$J(\theta) = \frac{1}{2N} \sum_{i=1}^{N} (y_i - h_\theta(x_i))^2$$

Gradient Descent

In order to find the optimal hypothesis that fits the data, we need to minimize the cost function with respect to θ.

$$min_{\theta_0,\theta_1} J(\theta_0, \theta_1)$$

This minimization is done with a technique called gradient descent.

Gradient Descent

The technique works as follows:

- Start with some values of θ_0, θ_1
- Keep changing θ_0, θ_1 to decrease $J(\theta)$
- Stop changing $J(\theta)$ when minimum is reached

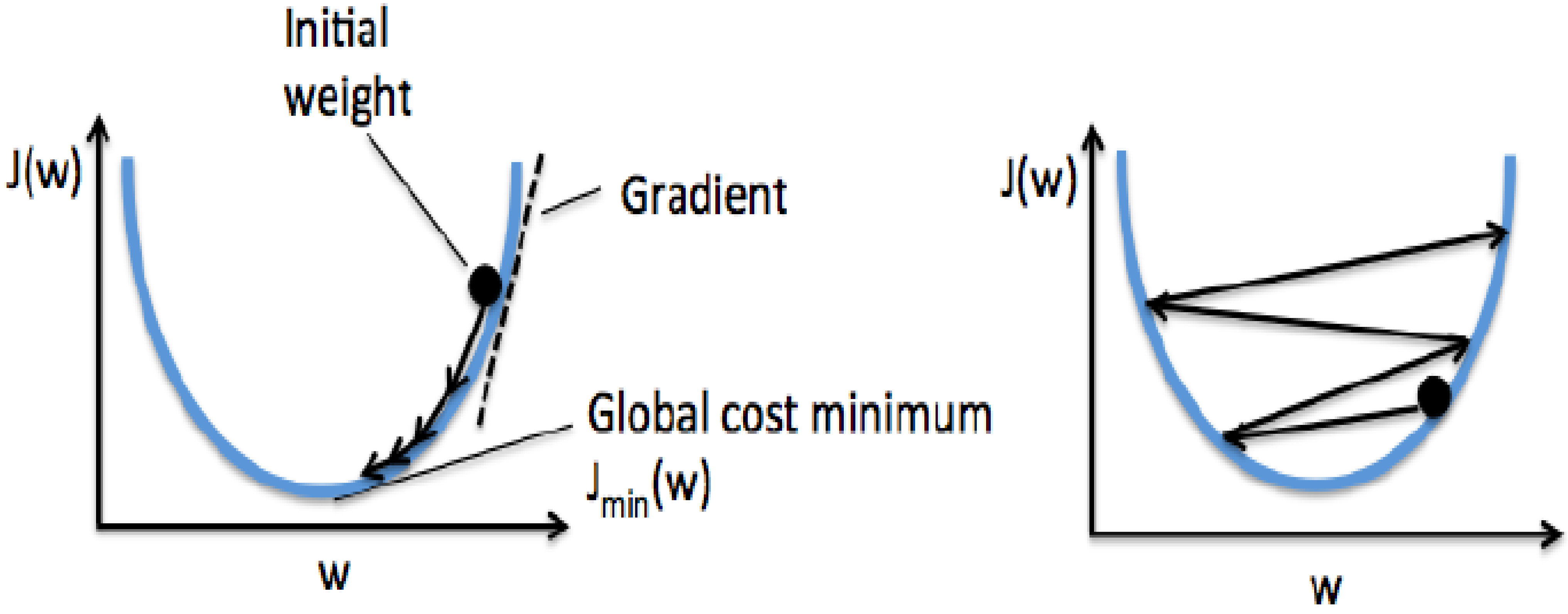
Gradient Descent

Algorithm

Repeat until convergence:

$$\theta_j := \theta_j - \underbrace{\alpha}_{\text{learning rate}} \times \underbrace{\frac{\partial J(\theta_j)}{\partial \theta_j}}_{\text{partial derivative along } \theta_j}$$

Gradient Descent intuition



Gradient Descent for Linear Regression

For linear regression, we can computer the derivative of the cost function to obtain the update we have to make:

$$\theta_j := \theta_j - \frac{\alpha}{N} \sum_{i=1}^{N} (h_\theta(x_i) - y_i) \times x_i^j$$

with $j = 0, 1,\ x_i^0 = 1$ and $x_i^1 = x$

Gradient Descent for Linear Regression

In order to accelerate (and sometimes even allow) the convergence of gradient descent, we should standarize the features as follows:

$$x_j' = \frac{(x_j - \mu_j)}{\sigma_j}$$

Important notions we did not cover

Here are some important notions we didn't cover in this session:

- How to avoid overfitting, or fitting data more than is warranted, using regularization (check ridge and lasso regression).

- How to evaluate a model using cross-validation.

Thank you for your attention!

Let’s go ahead and start the hands-on session!