Machine Learning

Chapter 2 지도 학습(Supervised Learning)

START

Smart Media
스마트미디어인재개발원

학습목표

- 선형회귀 모델을 이해하고 사용 할 수 있다.
- Mean Squared Error를 이해 할 수 있다.
- 경사하강법을 이해 할 수 있다.
- 회귀 모델의 평가방법을 알 수 있다.

Linear Model

(Regression)

Smart Media
스마트미디어인재개발원

Linear Model

Linear Model (선형 모델)

- 입력 특성에 대한 선형 함수를 만들어 예측을 수행
- 다양한 선형 모델이 존재한다
- 분류와 회귀에 모두 사용 가능

Linear Model - Regression

회귀의 선형모델

x(hour)	y(score)
9	90
8	80
4	40
2	20

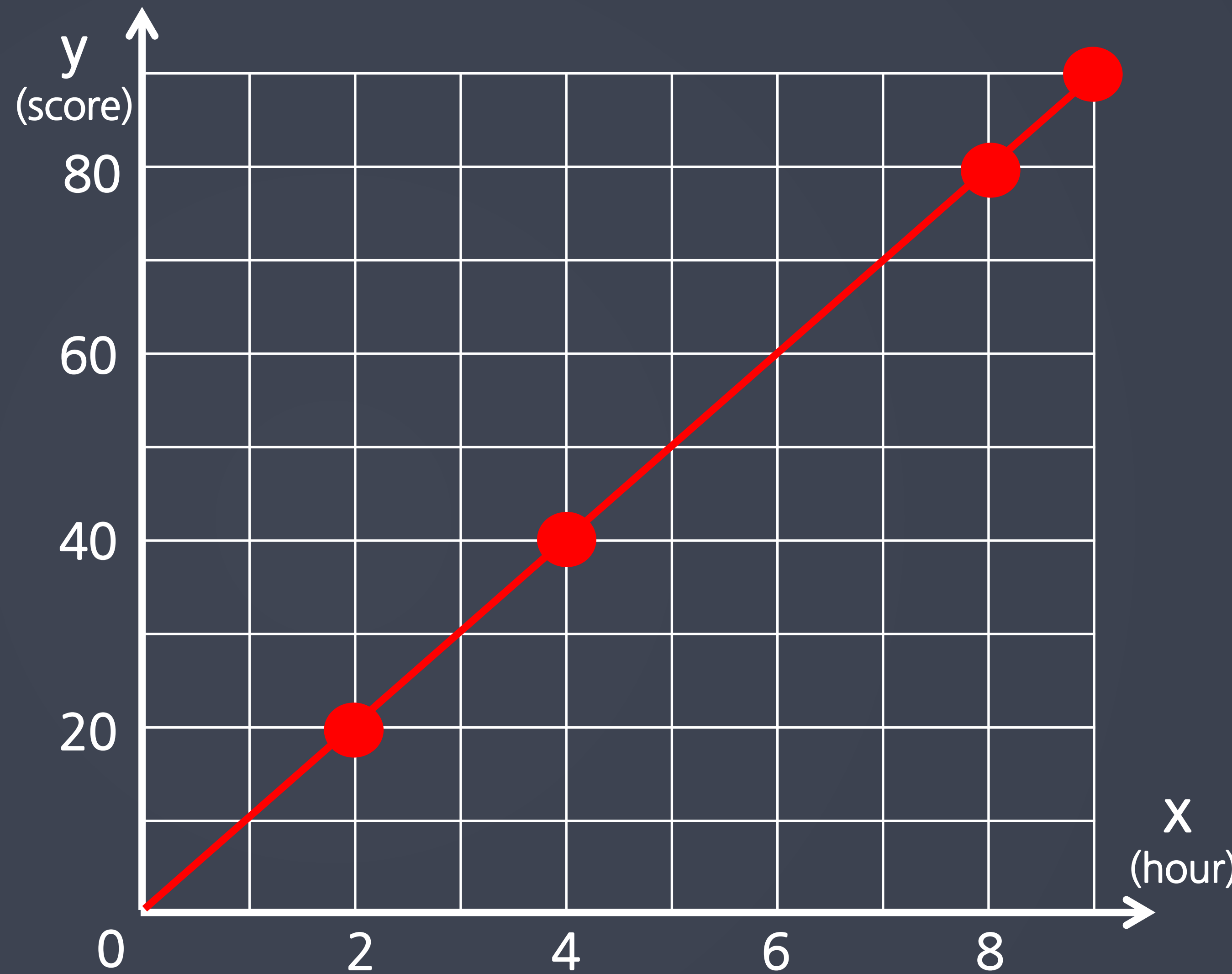
시험성적 데이터

7시간 공부 할 경우
성적은 몇 점 일까?

Linear Model - Regression

시험 성적 데이터

x(hour)	y(score)
9	90
8	80
4	40
2	20

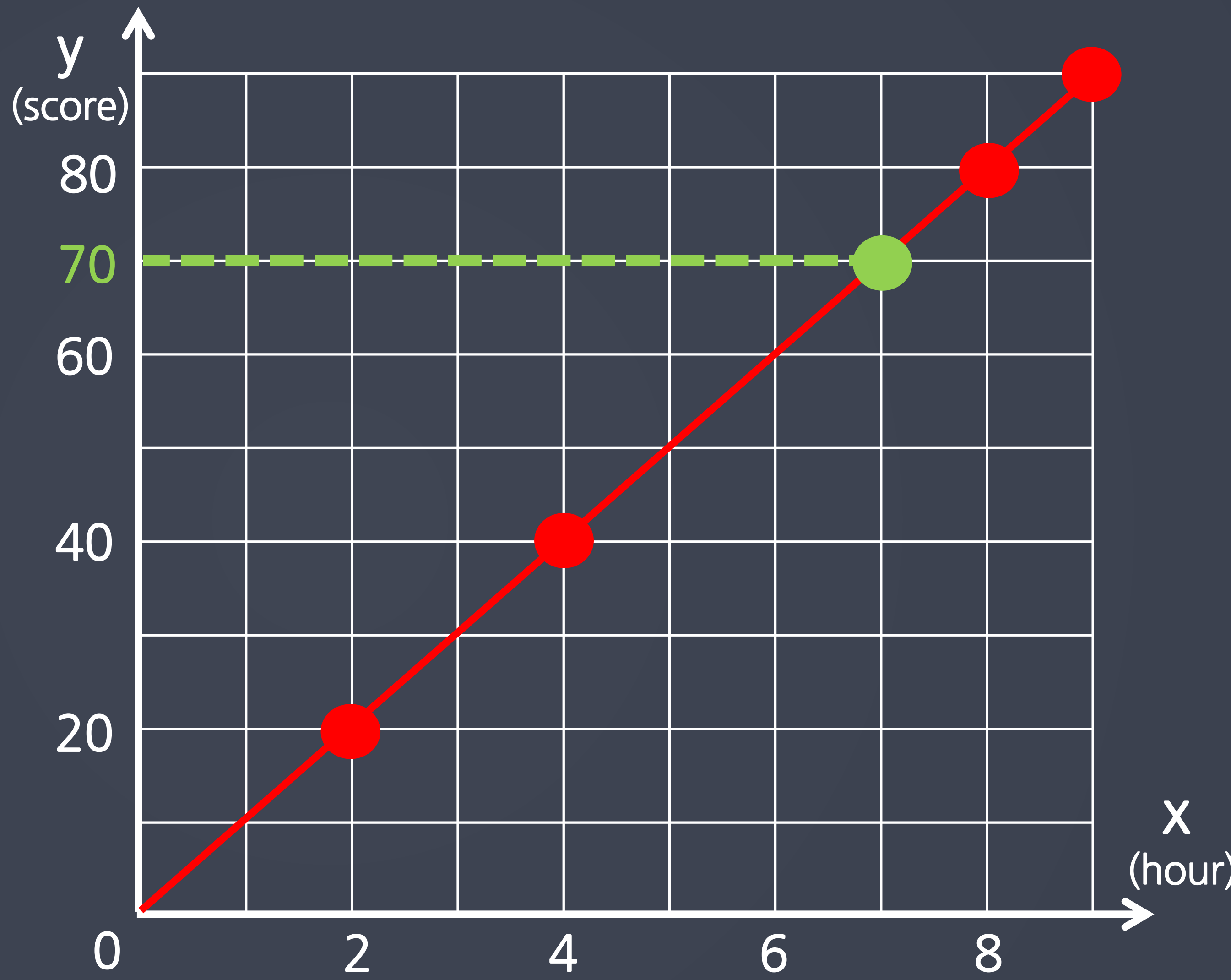


Linear Model - Regression

시험 성적 데이터

$$y = ax + b$$

$$y = 10x + 0$$



Linear Model - Regression

$$y = \underbrace{a}_{\text{기울기}}x + \underbrace{b}_{\text{절편}}$$

Linear Model - Regression (MSE)

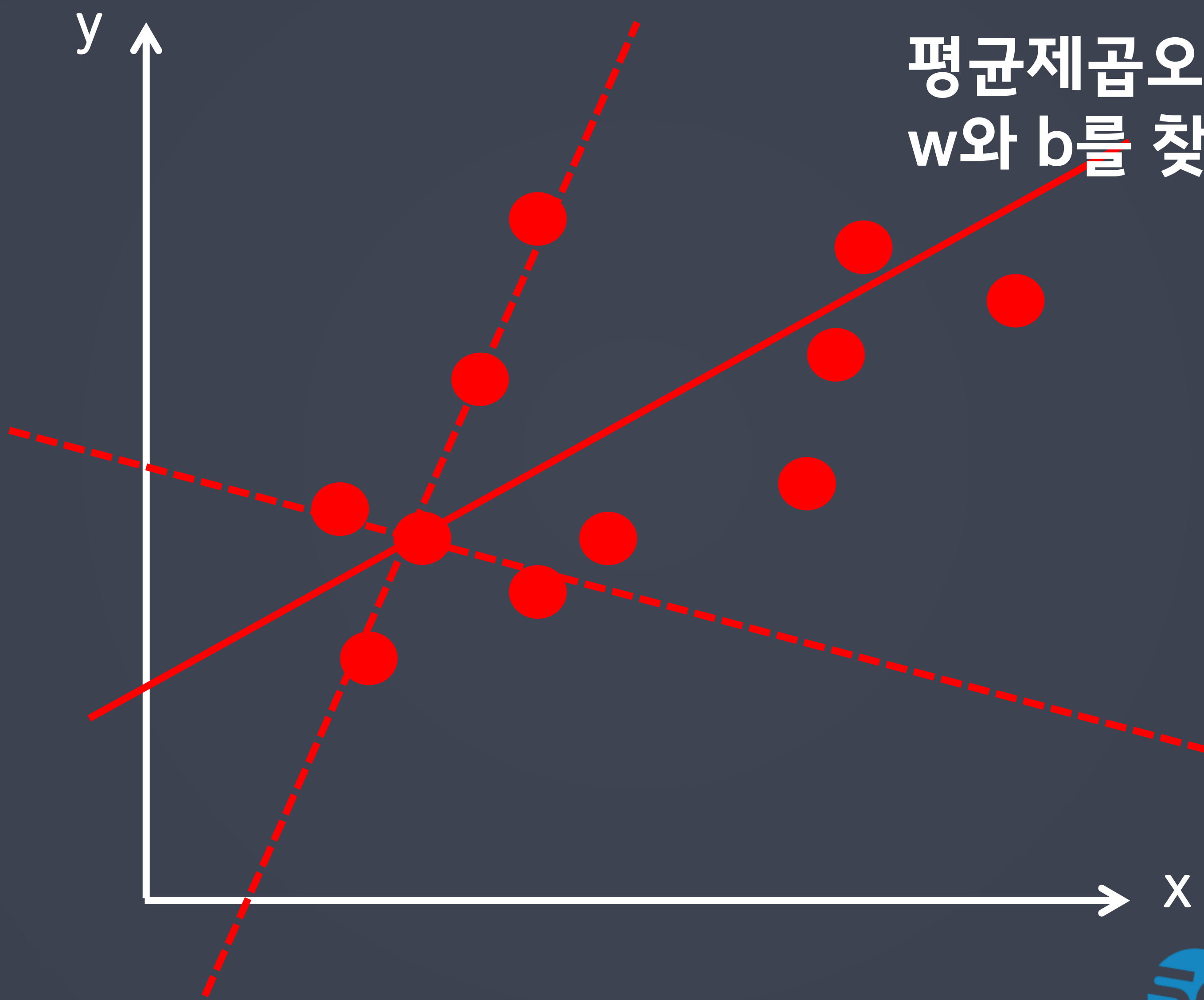
선형 회귀 함수

$$\hat{y} = w_1x_1 + w_2x_2 + w_3x_3 + \cdots + w_px_p + b$$

- w : 가중치(weight), 계수(coefficient)
- b : 절편(intercept), 편향(bias)

- 모델 w 파라미터 : model.coef_
- 모델 b 파라미터 : model.intercept_

Linear Model - Regression (MSE)



평균제곱오차가 최소인 w와 b를 찾는다

Linear Model - Regression (MSE)

Cost function

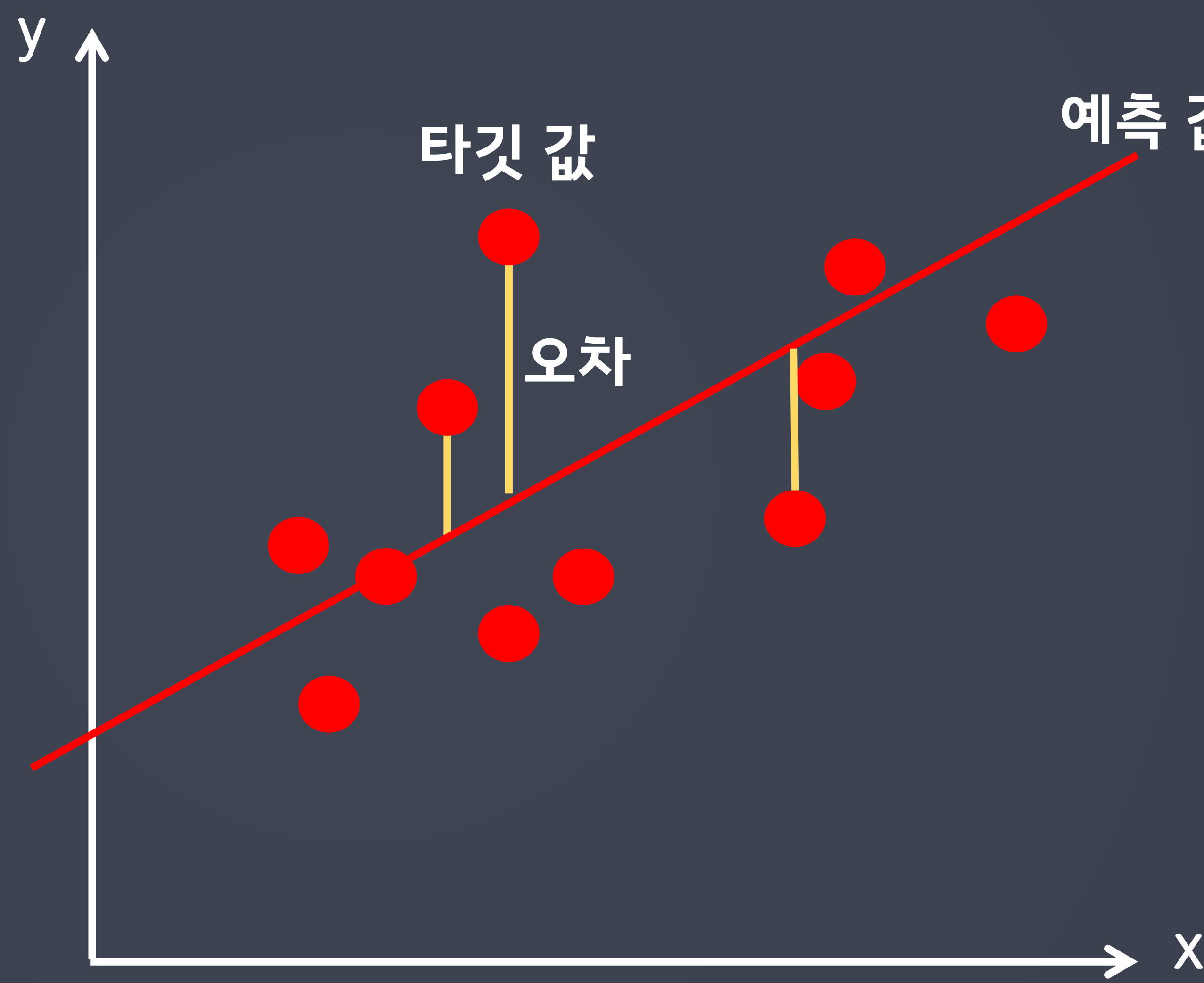
비용함수 : 수식을 검증

$H(x) = w * x + b$

$H(x) - y$

⬆

제곱 값 사용



Linear Model - Regression (MSE)

평균제곱오차 (Mean Squared Error)

RMSE를 사용하기도 한다

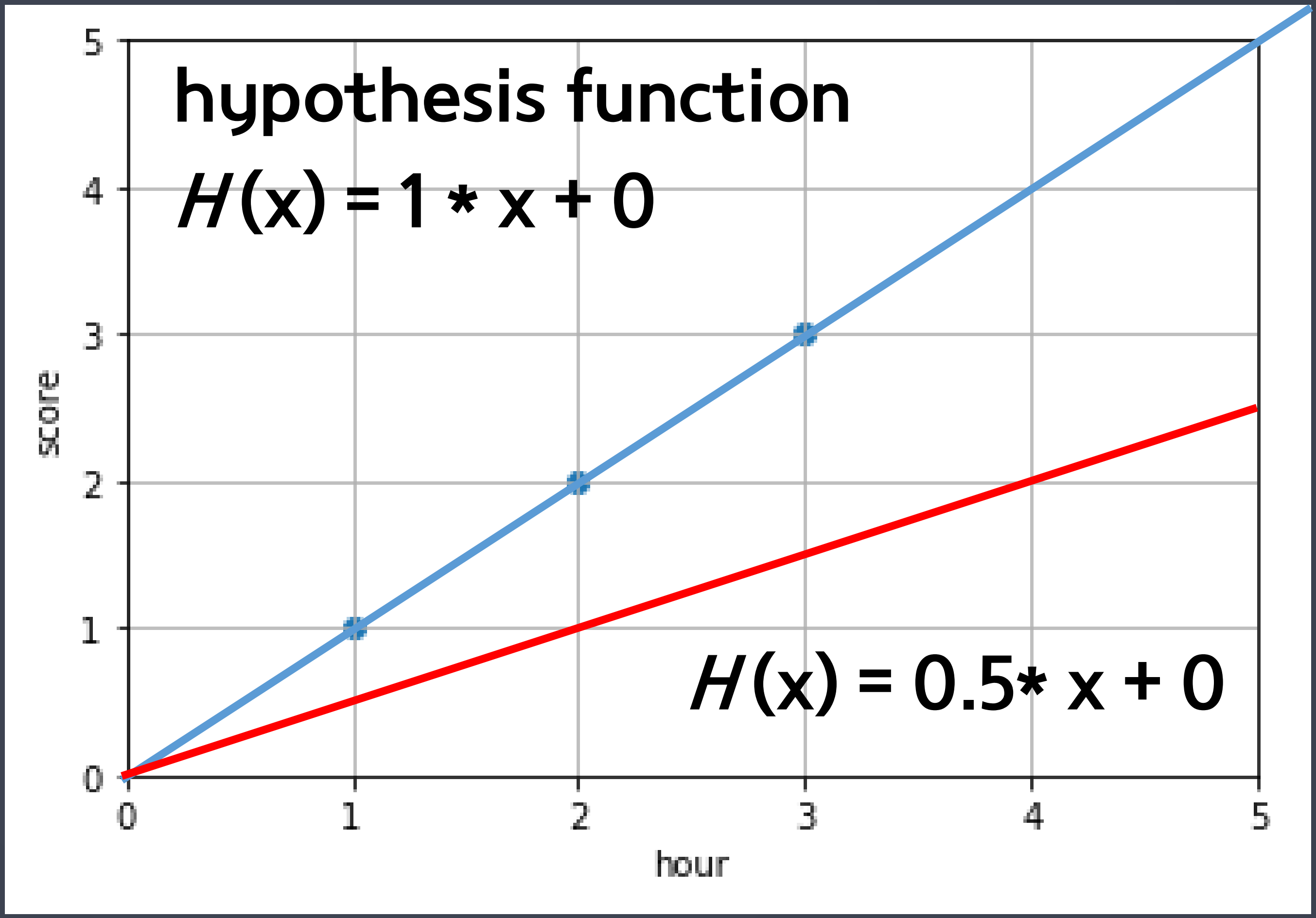
$$cost = \frac{1}{m}\sum_{i=1}^{m}(H(x_i) - y_i)^2$$

$$H(x) = Wx + b$$

Linear Model - Regression (MSE 값 계산)

두 가설의 MSE 값을 계산해보자.

x(hour)	y(score)
1	1
2	2
3	3



Linear Model - Regression

평균제곱오차(MSE)가 최소가 되는 w와 b를 찾는 방법

1. 수학 공식을 이용한 해석적 방법 (Ordinary Least Squares)
2. 경사하강법 (Gradient Descent Algorithm)

Linear Model - Regression

수학 공식을 이용한 해석적 방법 (Ordinary Least Squares)

$$a\sum x^2 + b\sum x = \sum xy$$

$$a\sum x + bn = \sum y$$

$$a = \frac{n\Sigma XY - \Sigma X \Sigma Y}{n\Sigma X^2 - \Sigma X \Sigma X}$$

$$b = \frac{\Sigma X^2 \Sigma Y - \Sigma X \Sigma XY}{n\Sigma X^2 - \Sigma X \Sigma X}$$

LinearRegression 클래스로 구현되어 있다.

Linear Model - Regression

수학 공식을 이용한 해석적 방법 (Ordinary Least Squares)

x(hour)	y(score)
1	1
2	2
3	3

$$a = \frac{n\Sigma XY - \Sigma X \Sigma Y}{n\Sigma X^2 - \Sigma X \Sigma X}$$

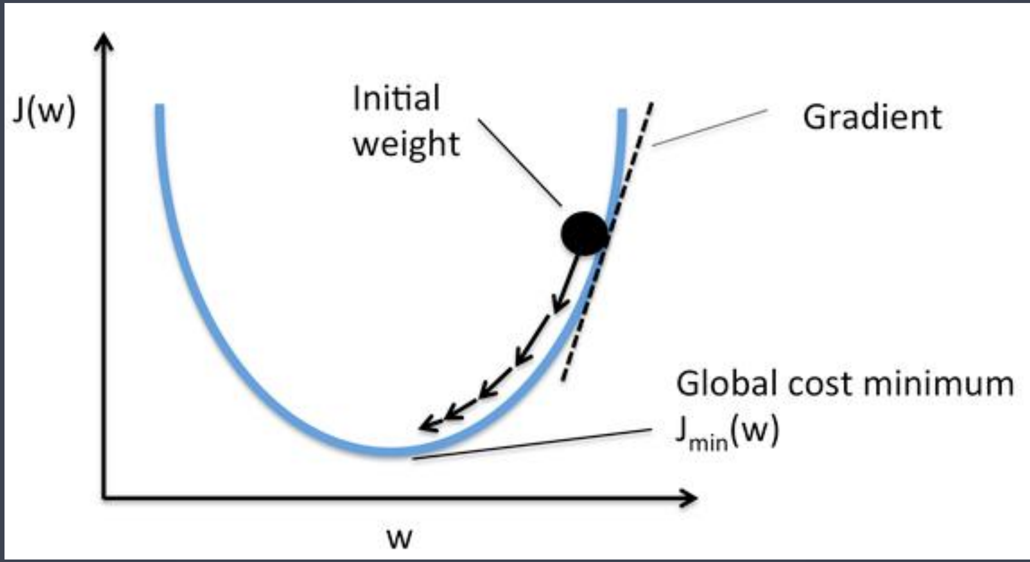
$$b = \frac{\Sigma X^2 \Sigma Y - \Sigma X \Sigma XY}{n\Sigma X^2 - \Sigma X \Sigma X}$$

LinearRegression 클래스로 구현되어 있다.

경사하강법 (Gradient Descent Algorithm)

Linear Model - Regression(Gradient descent algorithm)

경사하강법 (Gradient Descent Algorithm)



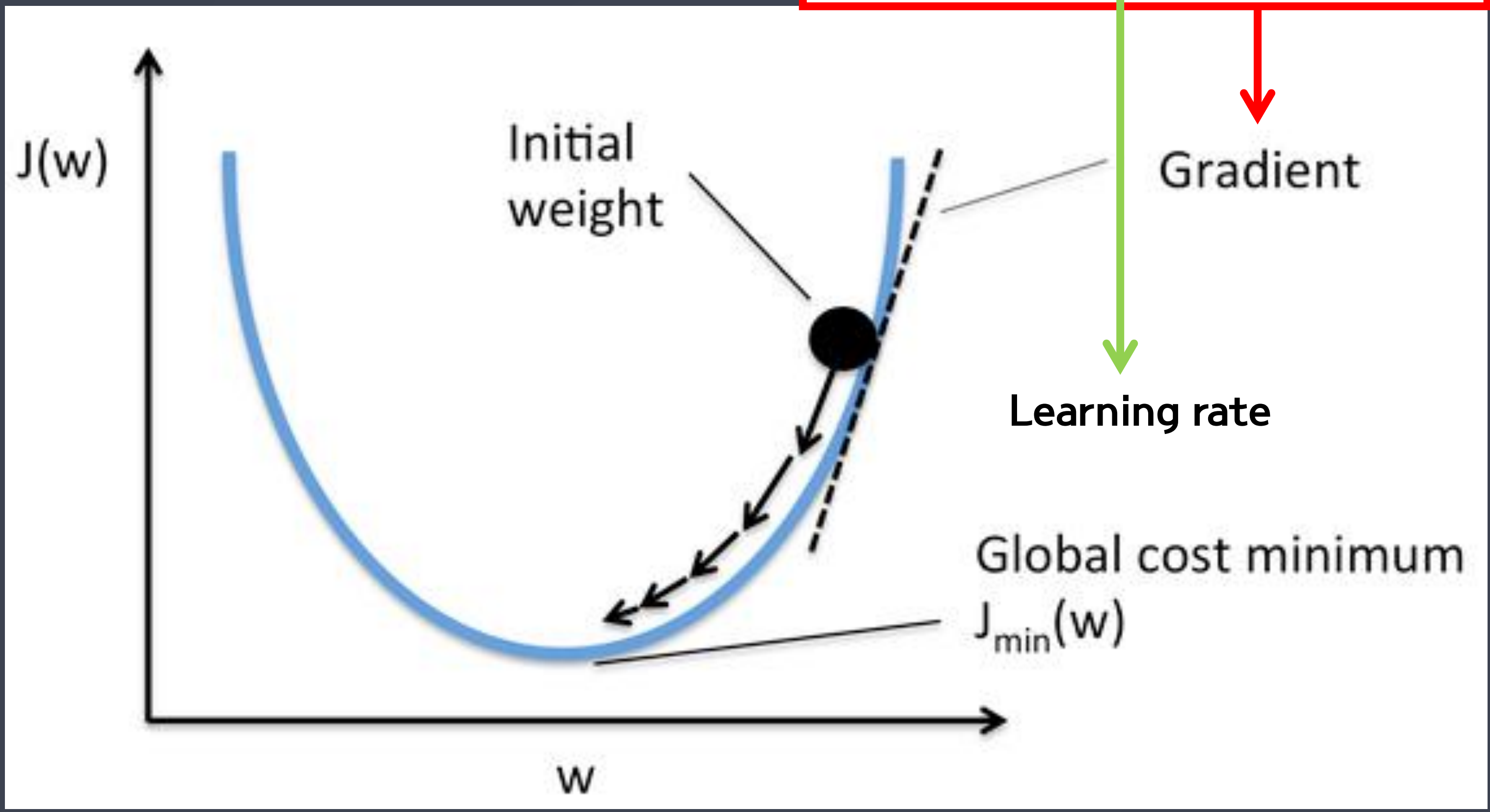
비용함수의 기울기(경사)를 구하여 기울기가 낮은 쪽으로
계속 이동하여 값을 최적화 시키는 방법

Linear Model - Regression(Gradient descent algorithm)

$$W := W - \boxed{\alpha} \frac{\partial}{\partial W} cost(W)$$

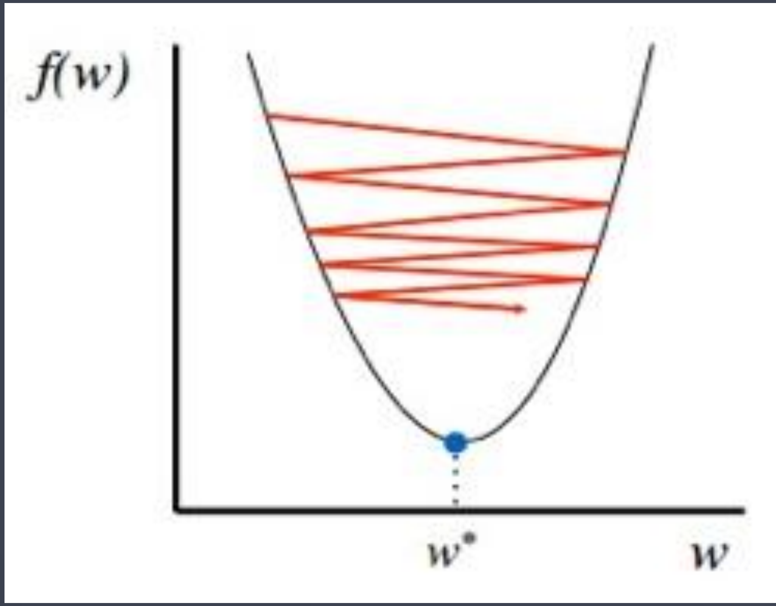
Gradient

Learning rate

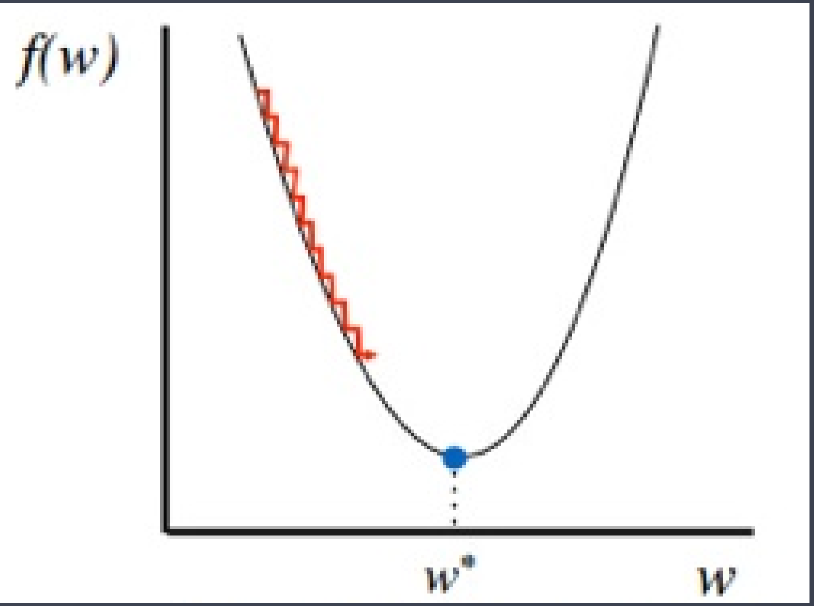


Linear Model - Regression

Learning rate가 큰 경우



Learning rate가 작은 경우



Linear Model - Regression

Linear Model 장점

- 결과예측(추론) 속도가 빠르다.
- 대용량 데이터에도 충분히 활용 가능하다.
- 특성이 많은 데이터 세트라면 훌륭한 성능을 낼 수 있다.

Linear Model - Regression

Linear Model 단점

- 특성이 적은 저차원 데이터에서는 다른 모델의 일반화 성능이 더 좋을 수 있다. ➡ 특성확장을 하기도 한다.

- LinearRegression Model은 복잡도를 제어할 방법이 없어 과대적합 되기 쉽다.

⬇

모델 정규화(Regularization)을 통해 과대적합을 제어한다.

Linear Model - Regularization

정규화

- 가중치(w)의 값을 조정하여 제약을 주는 것.

- L1 규제 : Lasso

  w의 모든 원소에 똑같은 힘으로 규제를 적용하는 방법. 특정 계수들은 0이 됨.
  특성선택(Feature Selection)이 자동으로 이루어진다.

- L2 규제 : Ridge

  w의 모든 원소에 골고루 규제를 적용하여 0에 가깝게 만든다.

Linear Model - Regularization

정규화 : cost 함수

alpha hpyerparameter로 조정

L1 규제 : Lasso

$$J(w)_{LASSO} = \sum_{i=1}^{n}(y^{(i)} - y^{\wedge(i)})^2 + \lambda \sum_{j=1}^{m}|w_j|$$

L2 규제 : Ridge

$$J(w)_{Ridge} = \sum_{i=1}^{n}(y^{(i)} - y^{\wedge(i)})^2 + \lambda \sum_{j=1}^{m}{w_j}^2$$

Linear Model - Regression 평가지표

- 회귀에서는 손실함수로 MSE를 주로 사용한다.
- 그런데 MSE 값은 성능 지표로 사용하기에는 부족하다.
  - 예를 들어 키를 예측하였는데 MSE 값이 5.7cm이 나왔다고 하면 이것의 성능이 얼마나 우수한지 다른 경우와 비교하기 어렵다.
  - 몸무게를 예측하였는데 MSE값이 3.8kg이라면 얼마나 우수한 것인가?
- 회귀분석에서는 **Coefficient of Determination (R-squared)**를 주로 사용한다

$$R^2 \equiv 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}$$

$$SS_{\text{res}} = \sum_i (y_i - f_i)^2 = \sum_i e_i^2$$

$$SS_{\text{tot}} = \sum_i (y_i - \bar{y})^2$$