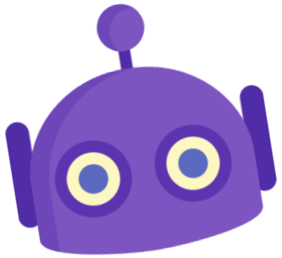
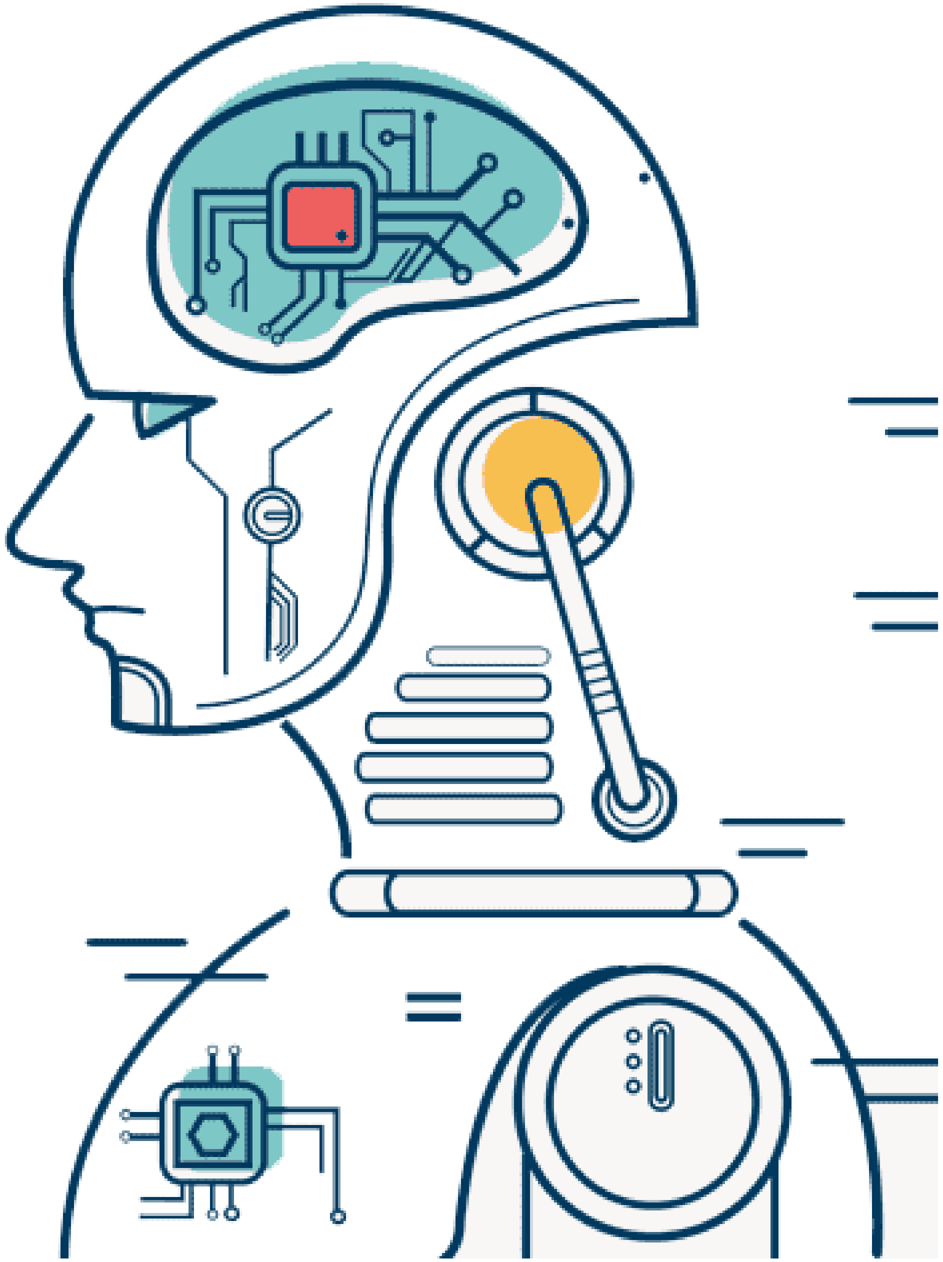
Smart Media
스마트미디어인재개발원



Linear Model
(Regression)



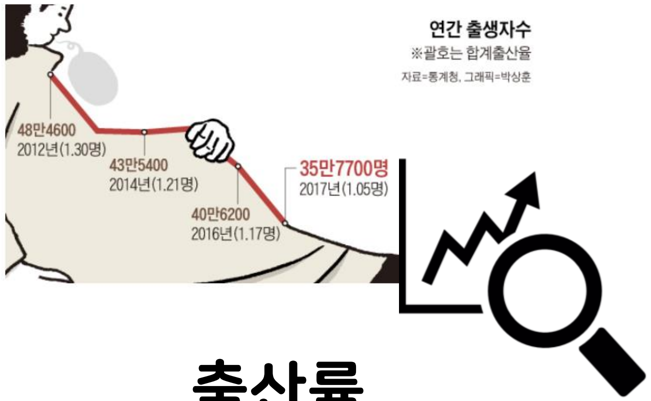
Linear Model - Regression

회귀 (Regression)

- 분류보다 회귀를 많이 사용.
- 오랜 기간 동안 현대 통계학에서 발생되어 다양하게 이용되는 학문분야.

주식

집값



출산률

Linear Model (선형 모델)

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

Linear Model (선형 모델)

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

시험성적 데이터

7시간 공부 할 경우
성적은 몇 점 일까?

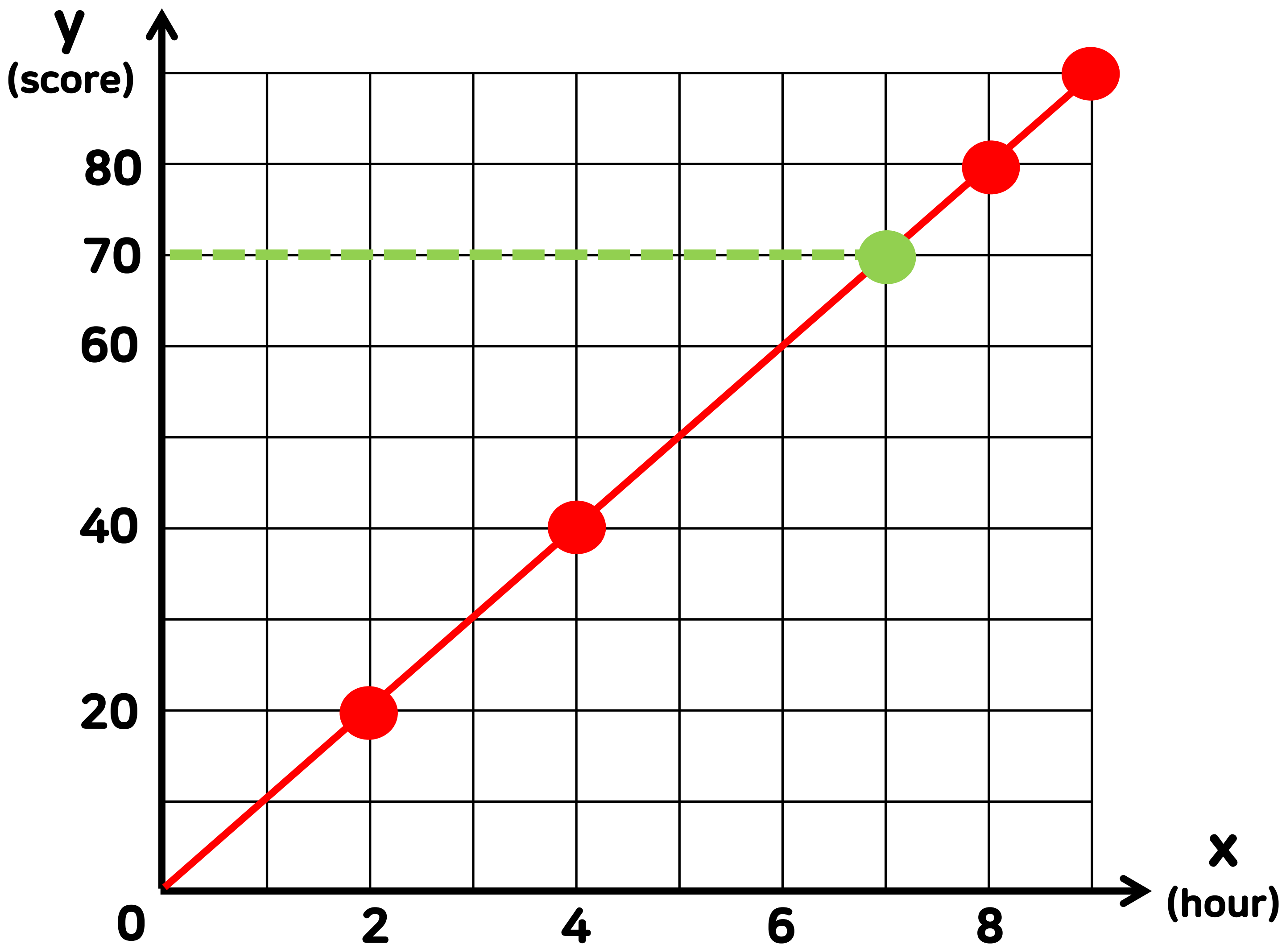
Linear Model - Regression

Smart Media
스마트미디어인재개발원

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

$$y = ax + b$$

$$y = 10x + 0$$



Linear Model - Regression

Smart Media
스마트미디어인재개발원

기울기 절편

$$y = ax + b$$

Linear Model - Regression

Smart Media 스마트미디어인재개발원

가중치 편향

$$y = ax + b$$

선형 회귀 함수

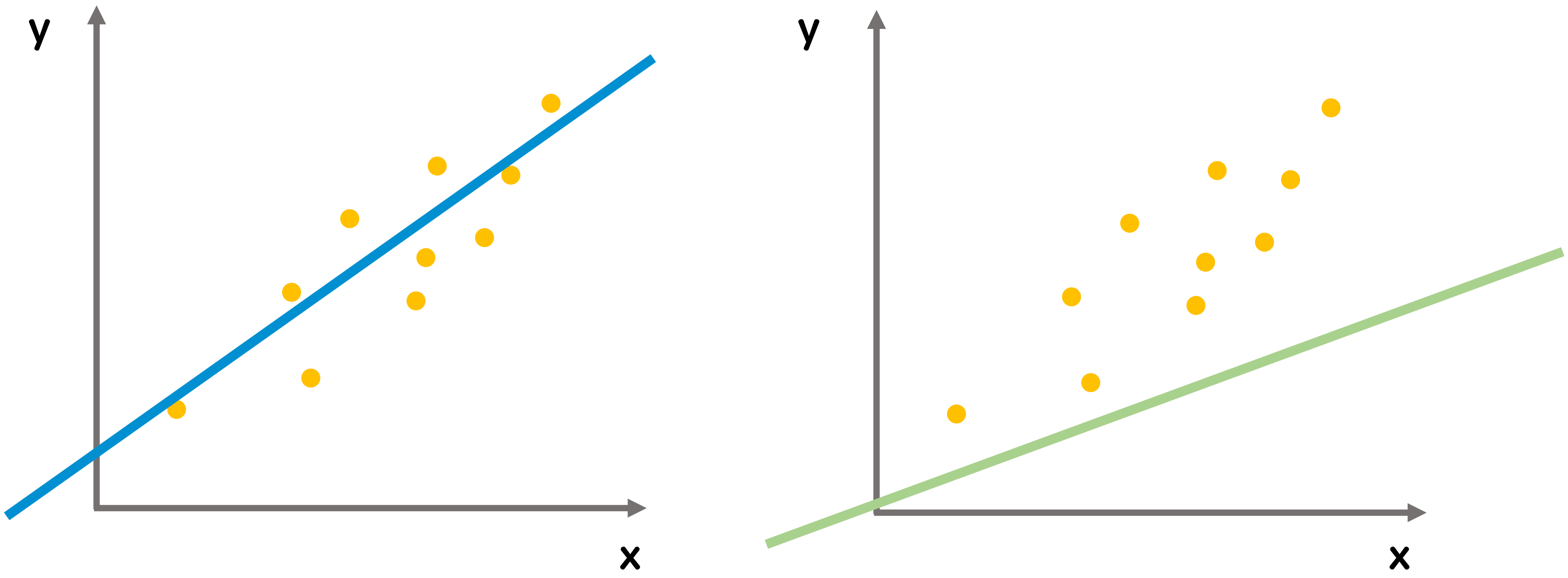
$$y = w_1x_1 + w_2x_2 + w_3x_3 + \cdots + w_px_p + b$$

- w : 가중치(weight), 계수(coefficient)
- b : 편향(bias), 절편(intercept)
-
- 모델 w 파라미터 : model.coef_
- 모델 b 파라미터 : model.intercept_

머신러닝 기법에서의 회귀

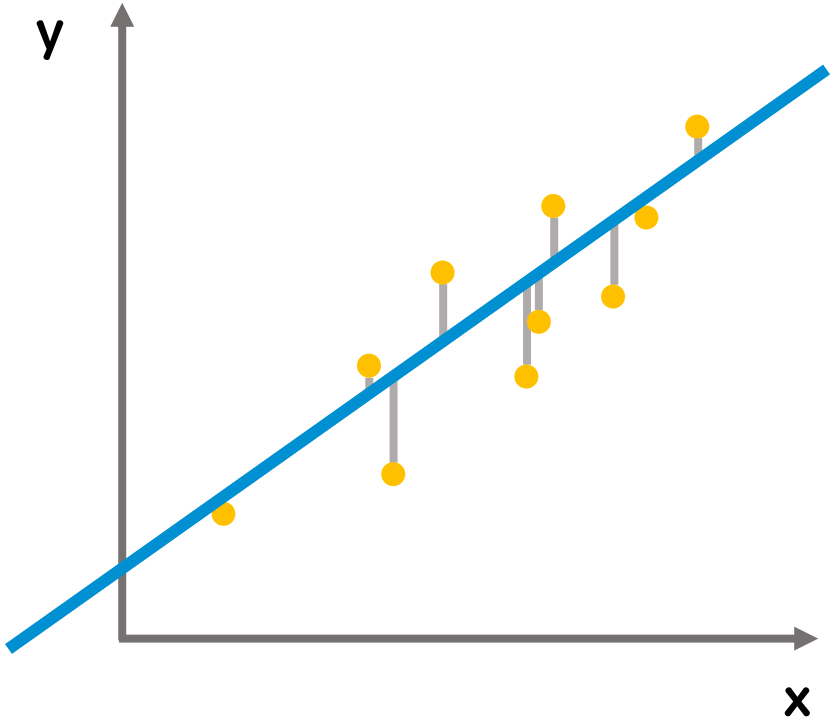
주어진 특성과 결정 값 데이터에 기반하여,
학습을 통해 최적의 **회귀 계수(W, b)**를 찾아내는 것.

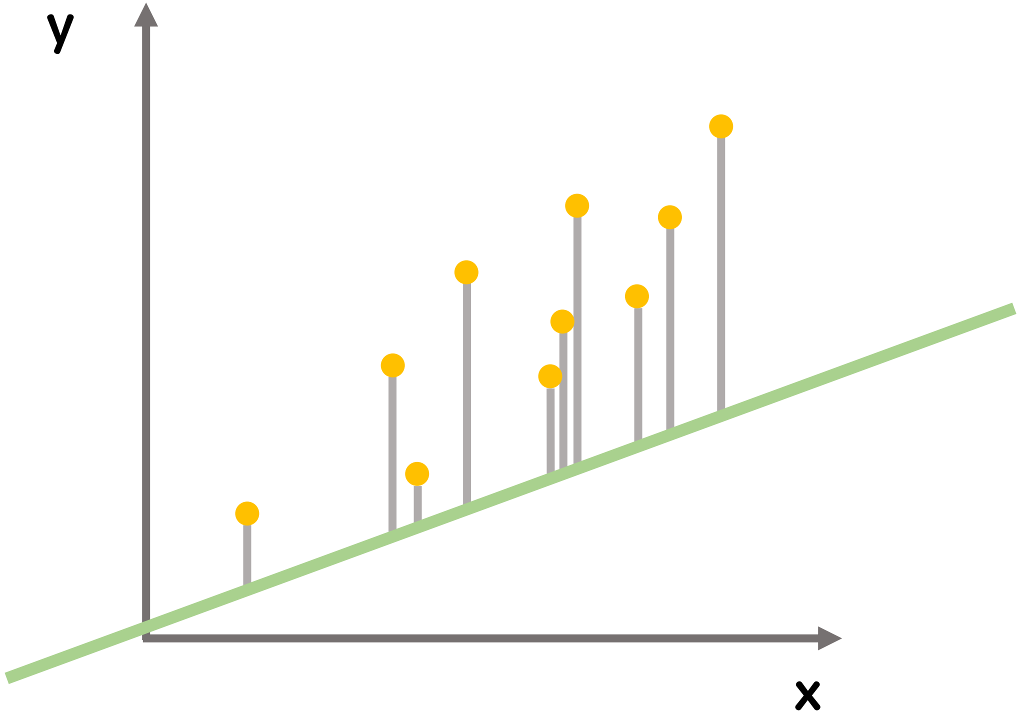
Linear Model - Regression



둘 중 어느 직선이 더 적합할까?

Linear Model - Regression





둘 중 어느 직선이 더 적합할까?

Linear Model – Regression

Cost function

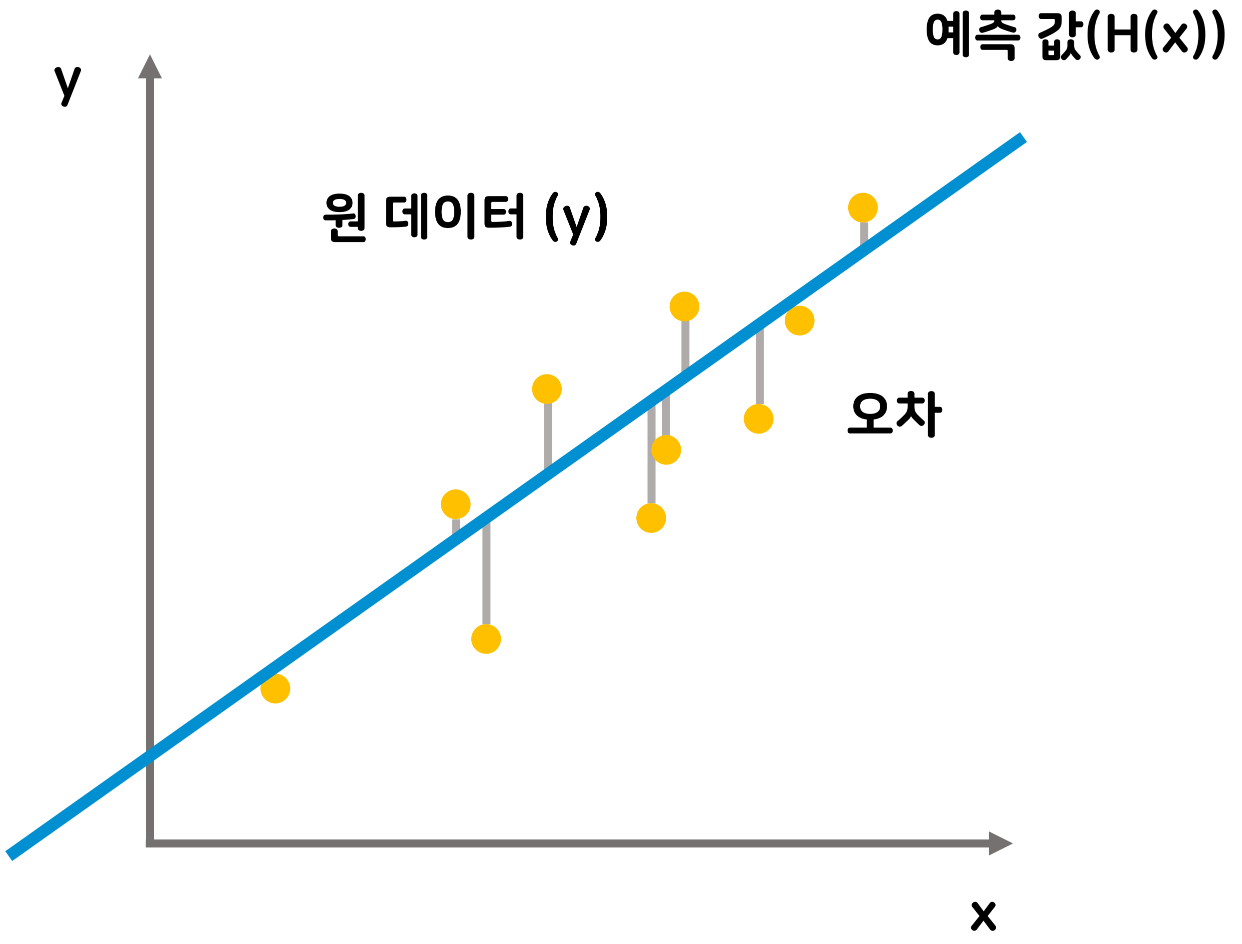
비용함수 : 수식을 검증

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

오차 = $H(x) - y$

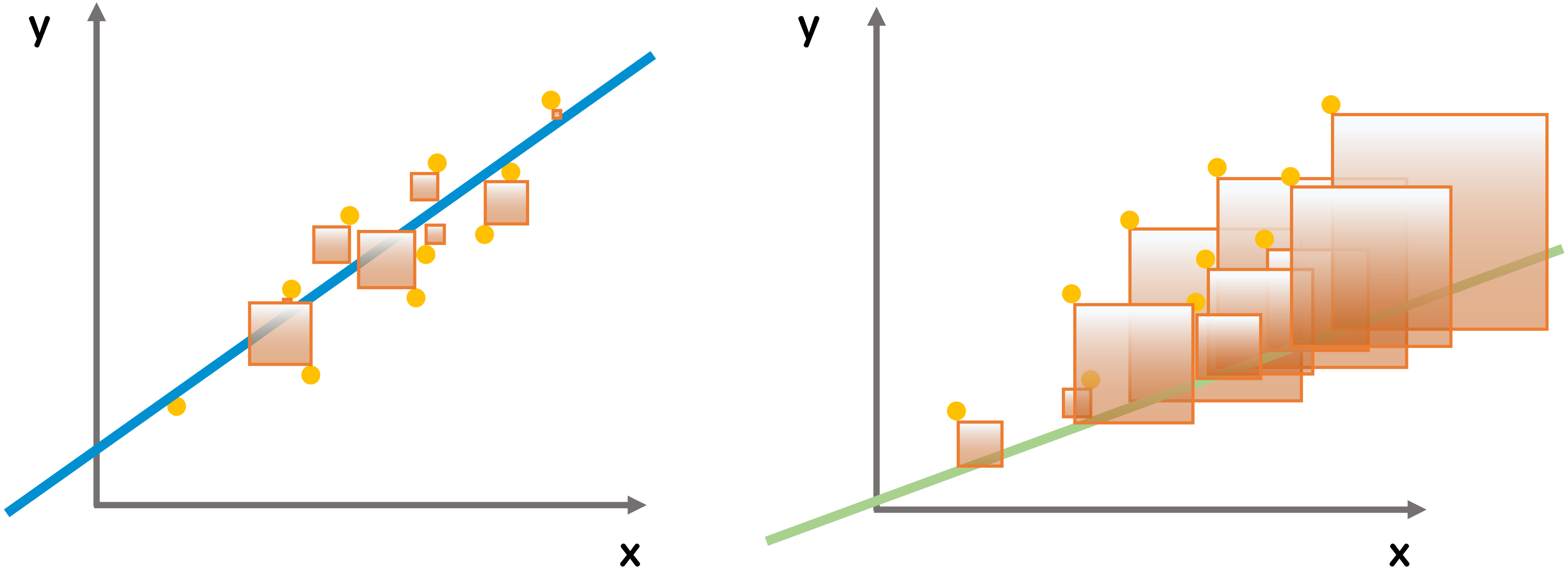
↑

제곱 값 사용



Linear Model - Regression

Smart Media
스마트미디어인재개발원



Linear Model - Regression

평균제곱오차 (MSE : Mean Squared Error)

$$MSE = \underset{\text{평균}}{\frac{1}{m}} \sum_{i=1}^{m} \underset{\text{오차}}{(H(x_i) - y_i)}^{2\ \text{제곱}}$$

평균제곱근오차 (RMSE : Root Mean Squared Error)

$$RMSE = \overset{\text{근}}{\sqrt{\underset{\text{평균}}{\frac{1}{m}} \sum_{i=1}^{m} \underset{\text{오차}}{(H(x_i) - y_i)}^{2\ \text{제곱}}}}$$

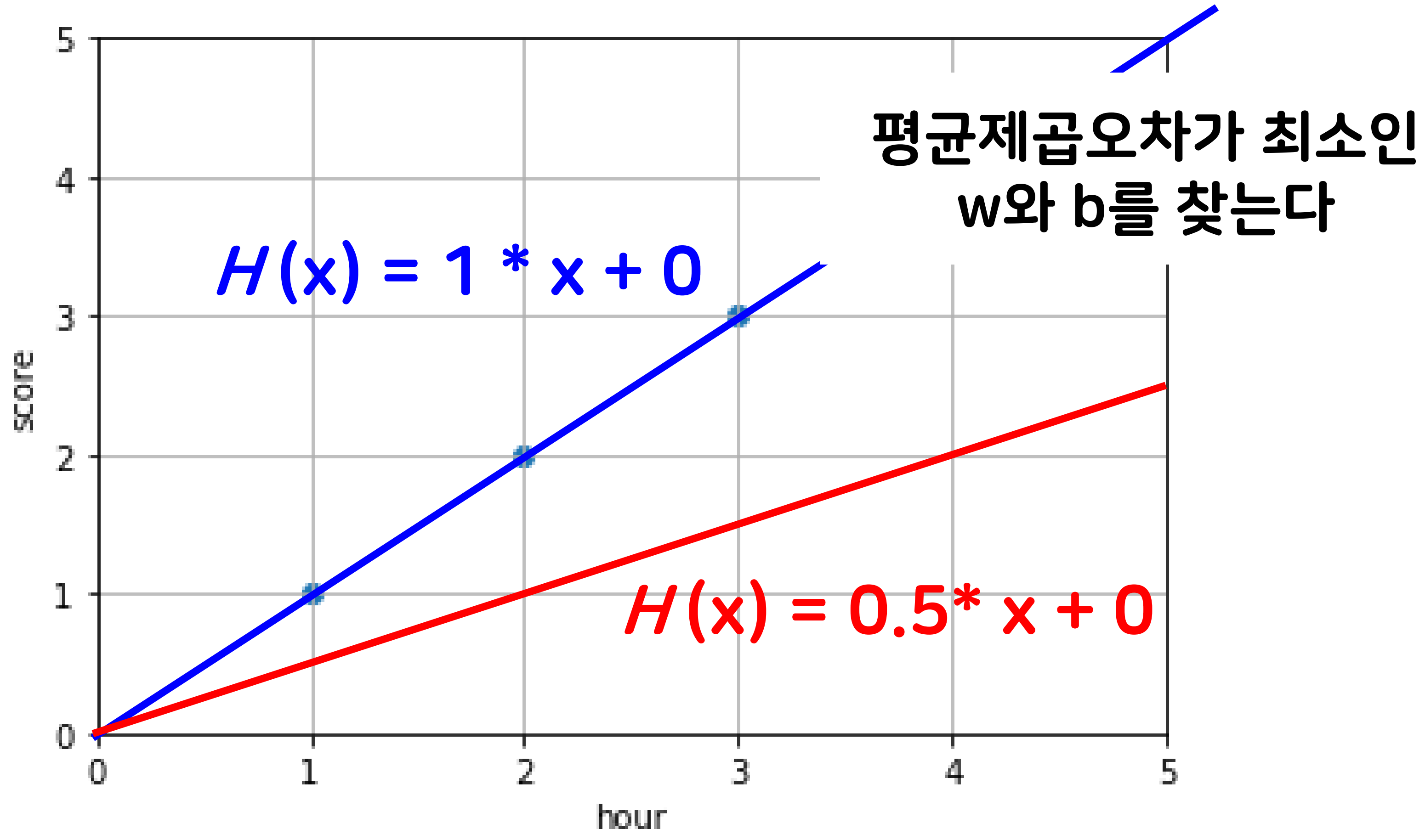
Linear Model – Regression (MSE)

Smart Media 스마트미디어인재개발원

두 데이터의 MSE 값을 계산해보자.

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

파란선 H(x) -> 0
빨간선 H(x) -> 1.167



Linear Model – Regression (MSE)

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

Smart Media 스마트미디어인재개발원

수학 공식을 이용한 해석적 방법 (Ordinary Least Squares)

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

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

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

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

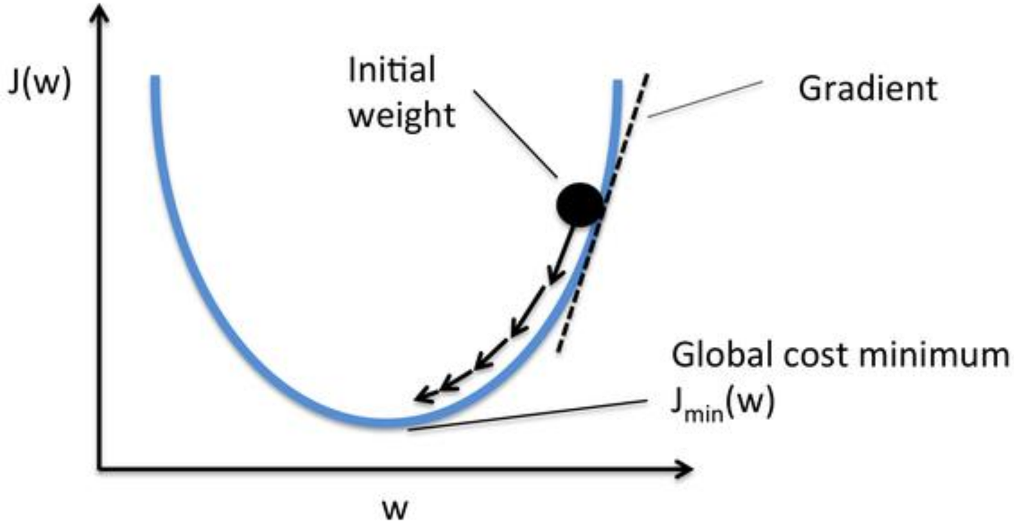
LinearRegression 클래스로 구현되어 있다.

LinearRegression 사용하기

경사하강법 (Gradient Descent Algorithm)

경사하강법 (Gradient Descent Algorithm)

경사하강법 (Gradient Descent Algorithm)

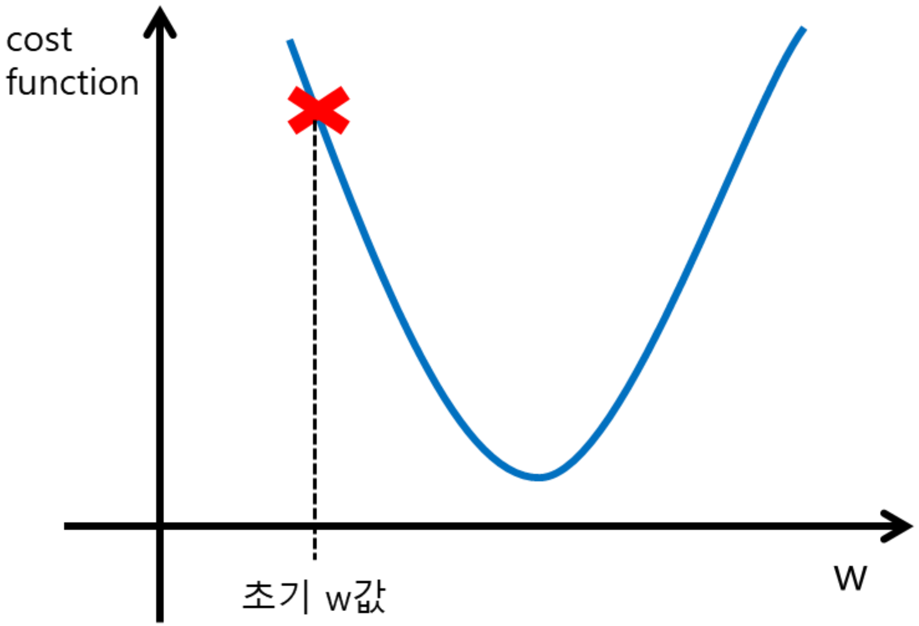


비용함수의 기울기(경사)를 구하여 기울기가 낮은 쪽으로 계속 이동하여 값을 최적화 시키는 방법

경사하강법 (Gradient Descent Algorithm)

우선 임의로 w값을 하나 선정

- 운이 아주 좋으면 최적의 값이겠지만 그렇지 않을 확률이 훨씬 더 큼
- 대부분 최적의 w값과는 거리가 먼 것이 설정



경사하강법 (Gradient Descent Algorithm)

(1) 최적의 w값을 찾아가기 위해서 시작점에서 손실 곡선의 기울기를 계산 → 비용함수를 w에 대해서 편미분

(2) 파라미터를 곱한 것을 초기 설정된 w값에서 빼 줌

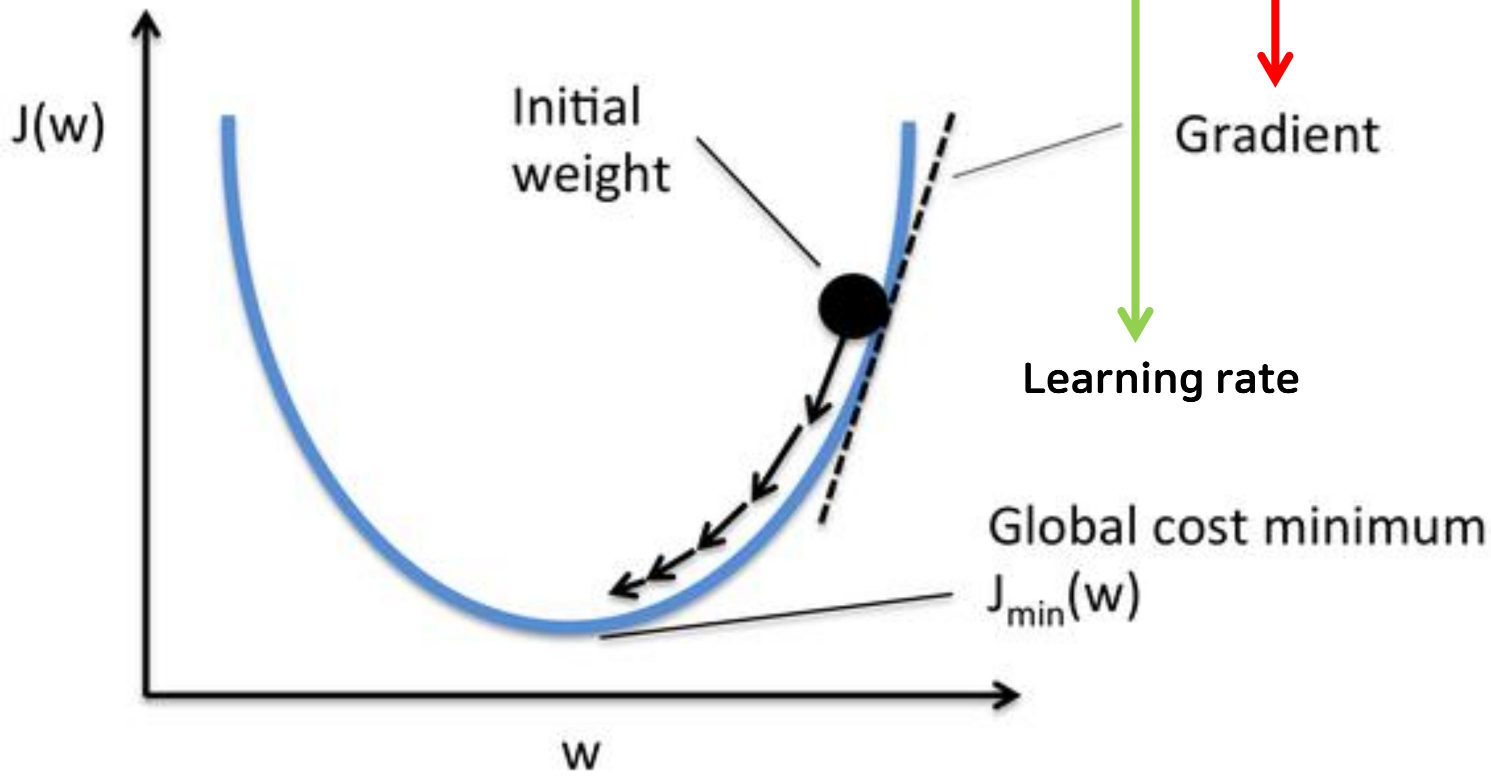
- **학습률(learning rate)** : 기울기의 보폭
- 학습률이 너무 작으면 최적의 w를 찾는데 오래 걸리고 크면 건너뛰어 버리고 발산할 수 있음

$$w' = w - \eta\frac{\partial e}{\partial w} \qquad b' = b - \eta\frac{\partial e}{\partial b}$$

Linear Model – Regression(Gradient descent algorithm)

경사하강법 (Gradient Descent Algorithm)

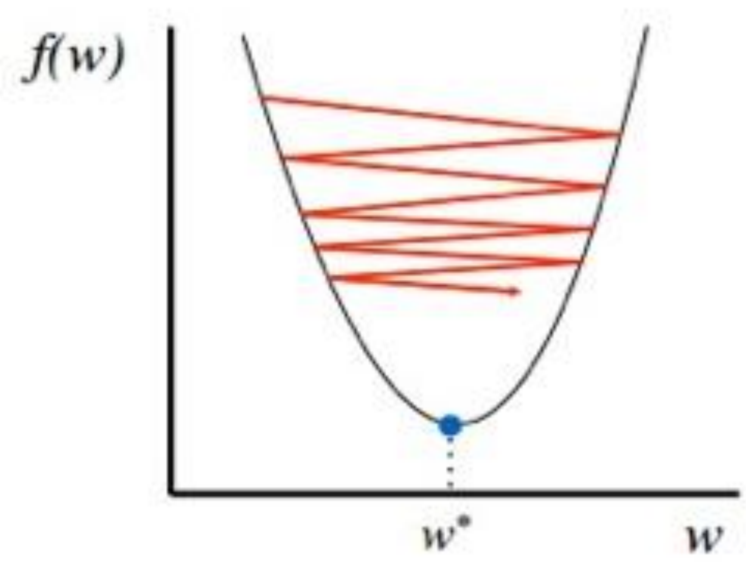
$$W := W - \boxed{\alpha} \frac{\partial}{\partial W} cost(W)$$



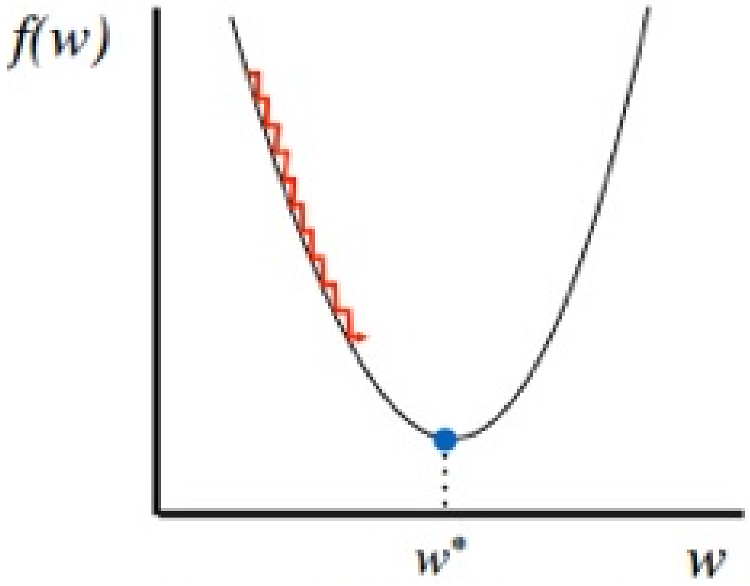
Smart Media 스마트미디어인재개발원

경사하강법 (Gradient Descent Algorithm)

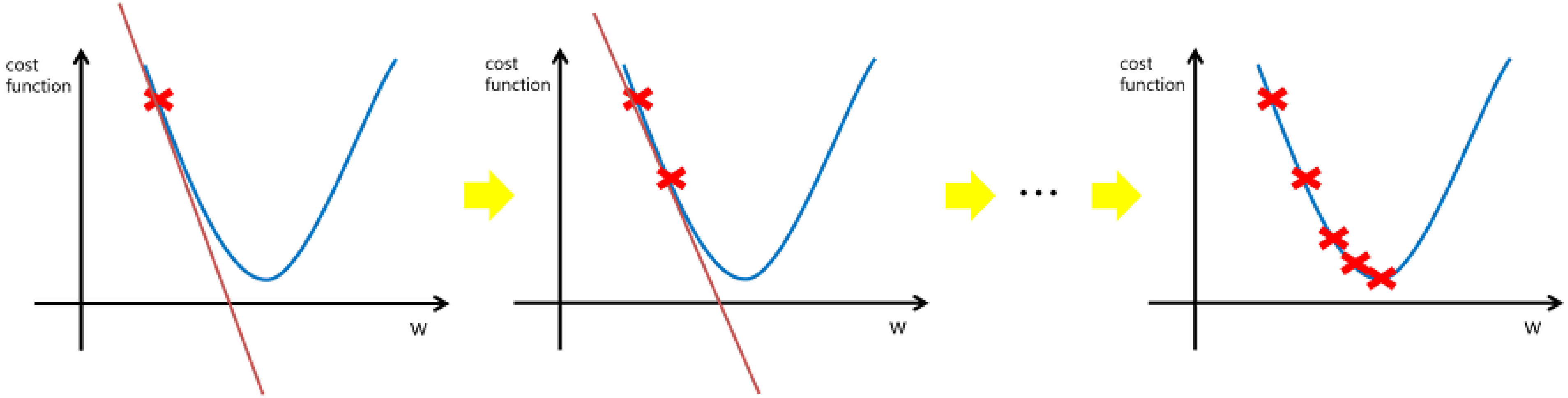
Learning rate가 큰 경우



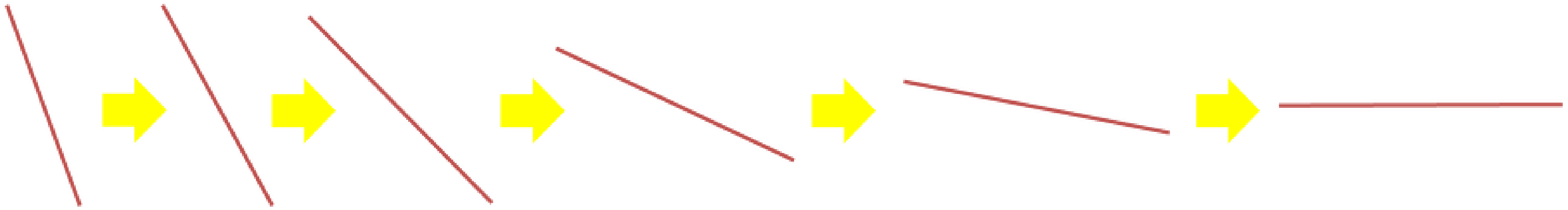
Learning rate가 작은 경우



경사하강법 (Gradient Descent Algorithm)



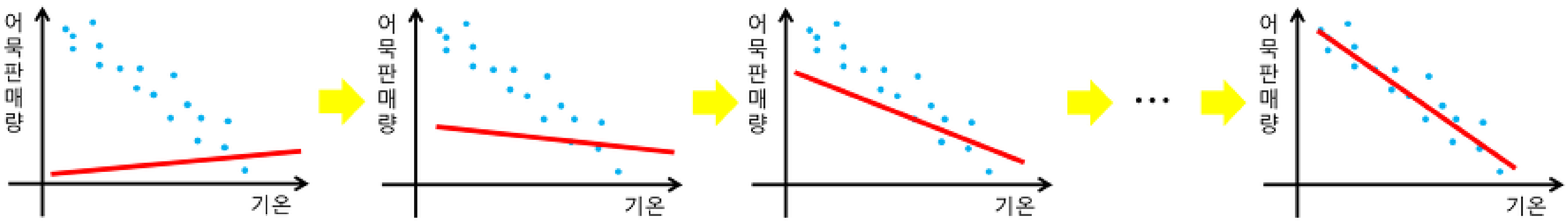
W값 갱신 과정



경사가 점차 감소되는 현상을 이용하므로 경사감소법!

Linear Model – Regression(Gradient descent algorithm)

경사하강법 (Gradient Descent Algorithm)



초기 w, b값은 데이터를 잘 반영하는 일차함수식을 만들지 못했지만,

점차적으로 데이터를 잘 반영해내는 값들로 갱신

주요 매개변수(Hyperparameter)

scikit-learn의 경우

```
SGDRegressor(max_iter, eta0)
```

- 가중치 업데이트 횟수 : max_iter
- 학습률 : eta0

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.