CONTENTS

02. Linear Regression (선형회귀 개념)

❑ 기계학습 예제를 위해 가장 간단한 형태의 데이터를 고려해보자

- 우리반 학생들의 수학 공부시간 대비 시험점수가 아래와 같다. 내가 7시간 공부하면 몇 점 받을 수 있을까?

이름	수학 공부시간	점수
야쓰오	2	3
오공	4	4
블리츠	6	5
페이커	8	6

02. Linear Regression (선형회귀 개념)

❑ 기계학습을 위해 모델링 예제로 바꿔보자

- 어떻게 모델링을 해야할까?

X = 연습시간 {2,4,6,8}

입력 x값에 **어떤 변환**을 하니까 y가 나온다

f(x) = ??

Y = 점수 {3,4,5,6}

02. Linear Regression (선형회귀 개념)

❑ 모델링 방법1: **선형회귀**

❑ 가설: 데이터의 관계를 선형으로 표현할 수 있지 않을까? (예, 공부시간 대비 성적은 비례)

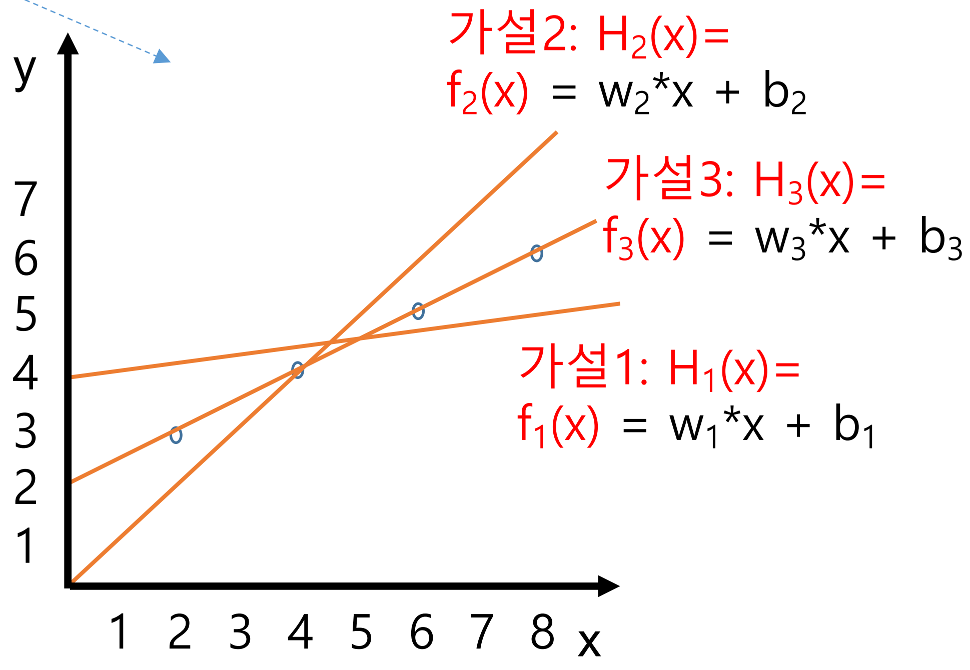
- 종속 변수 y와 한 개 이상의 독립 변수 X와의 **선형 상관 관계**를 모델링하는 회귀분석 기법 (위키)

X = 연습시간 {2,4,6,8}

입력 x값에 **선형 변환**을 하니까 y가 나온다

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

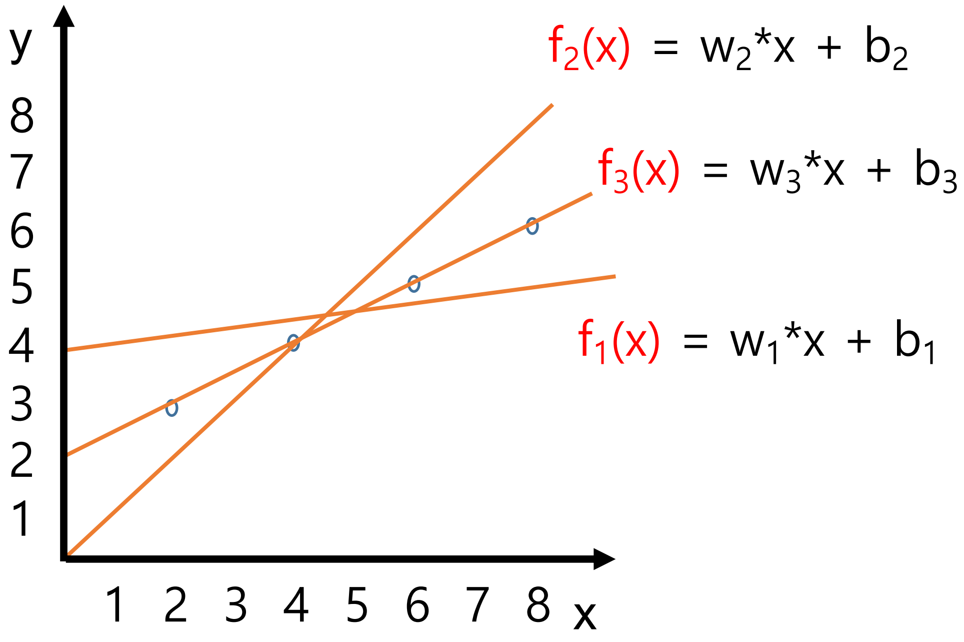
Y = 점수 {3,4,5,6}



02. Linear Regression (선형회귀 개념)

❑ 모델링 방법1: **선형회귀**

- 관측된 표본 데이터를 가장 잘 설명할 수 있는 선을 하나 찾는다. = 최적화된 w, b조합을 찾는다.

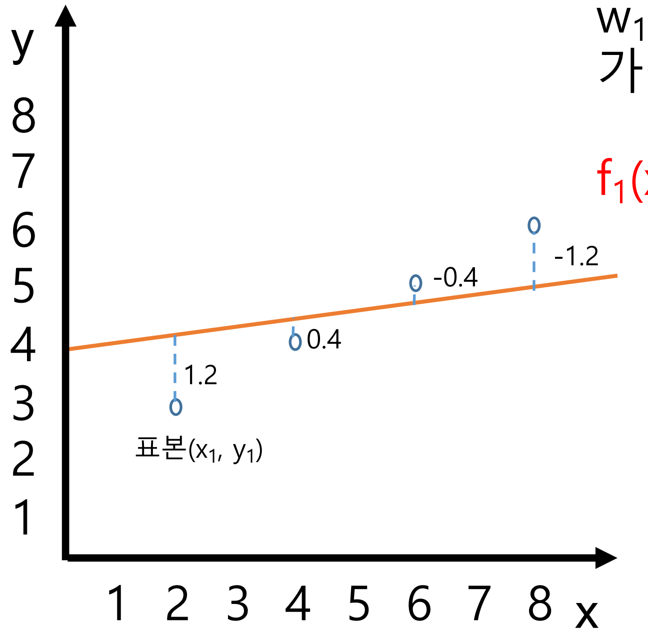


가능한 w와 b 조합 중에 어떤 모델이
다른 모델보다 더 좋다고 할 수 있을까?

평가할 척도가 필요하다!!

❑ 선형회귀 평가

- 관측된 표본데이터와 우리가 선택한 w, b값으로 생성한 직선을 비교해 차이가 적은 w, b값을 찾자



$w_1 = 0.1$, $b_1 = 0.4$로 가정해보자

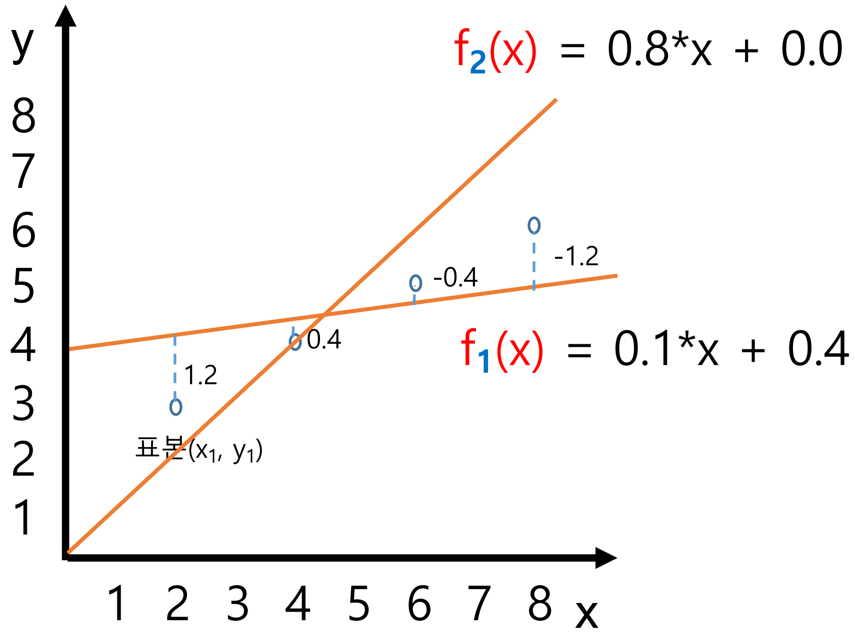
$f_1(x) = 0.1*x + 0.4$

어떤 평가척도를 사용해서 w, b를 평가할까?

- 관측된 표본(x_1, y_1)과 f1과의 **편차**= $(f_1(x_1) - y_1) = 1.2$
- 전체 편차의 합= $\Sigma_{i=1}(f_1(x_i) - y_i) = 0$??
- 전체 편차의 평균= $1/n*\Sigma_{i=1}(f_1(x_i) - y_i) = 0$??
- 편차 제곱의 평균= $1/n*\Sigma_{i=1}(f_1(x_i) - y_i)^2 = 0.8$

❑ 선형회귀 평가: Mean Squared Error(MSE) == 편차 제곱의 평균 (분산?)

- 과연 편차 제곱의 평균 (MSE)는 w, b에 따른 차이를 잘 반영할까?



$$cost(W,b) = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2$$

$f_1(x)$ = 0.1*x + 0.4

$$\mathbf{x}_1, y_1 \rightarrow (f_{\Theta_1}(2.0) - 3.0)^2 = ((0.1 * 2.0 + 4.0) - 3.0)^2 = 1.44$$
$$\mathbf{x}_2, y_2 \rightarrow (f_{\Theta_1}(4.0) - 4.0)^2 = ((0.1 * 4.0 + 4.0) - 4.0)^2 = 0.16$$
$$\mathbf{x}_3, y_3 \rightarrow (f_{\Theta_1}(6.0) - 5.0)^2 = ((0.1 * 6.0 + 4.0) - 5.0)^2 = 0.16$$
$$\mathbf{x}_4, y_4 \rightarrow (f_{\Theta_1}(8.0) - 6.0)^2 = ((0.1 * 8.0 + 4.0) - 6.0)^2 = 1.44$$

MSE=0.8

$f_2(x)$ = 0.8*x + 0.0

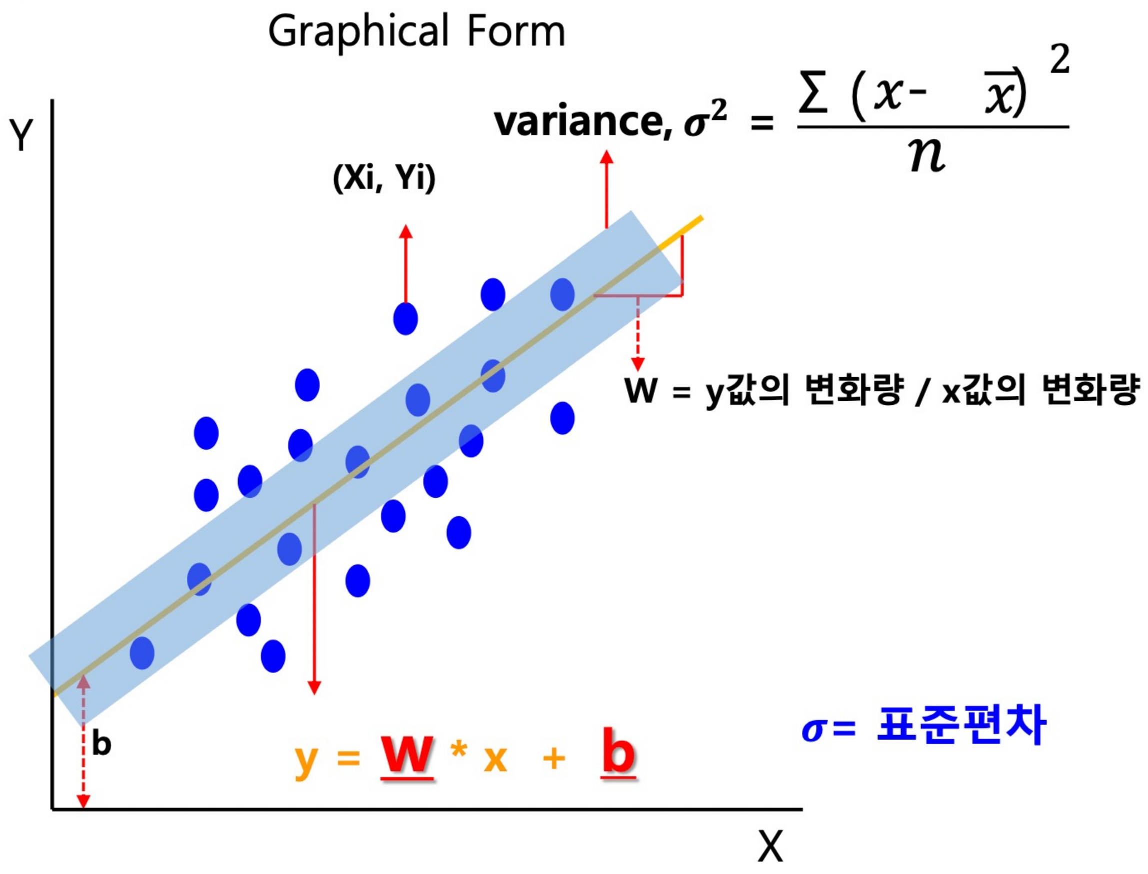
$$\mathbf{x}_1, y_1 \rightarrow (f_{\Theta_2}(2.0) - 3.0)^2 = ((0.8 * 2.0 + 0.0) - 3.0)^2 = 1.96$$
$$\mathbf{x}_2, y_2 \rightarrow (f_{\Theta_2}(4.0) - 4.0)^2 = ((0.8 * 4.0 + 0.0) - 4.0)^2 = 0.64$$
$$\mathbf{x}_3, y_3 \rightarrow (f_{\Theta_2}(6.0) - 5.0)^2 = ((0.8 * 6.0 + 0.0) - 5.0)^2 = 0.04$$
$$\mathbf{x}_4, y_4 \rightarrow (f_{\Theta_2}(8.0) - 6.0)^2 = ((0.8 * 8.0 + 0.0) - 6.0)^2 = 0.16$$

MSE=0.7

$f_3(x)$ = 0.5*x + 2 MSE=0

❑ 선형회귀 평가: Mean Squared Error(MSE)의 의미

- ❑ 분산은 평균을 기준으로 퍼져있는 정도를 나타내는 대푯값.
- ❑ 분산이 클 수록 데이터의 분포도가 평균에서 들쭉날쭉 불안정하다, 즉 **오차가 많다**는 의미

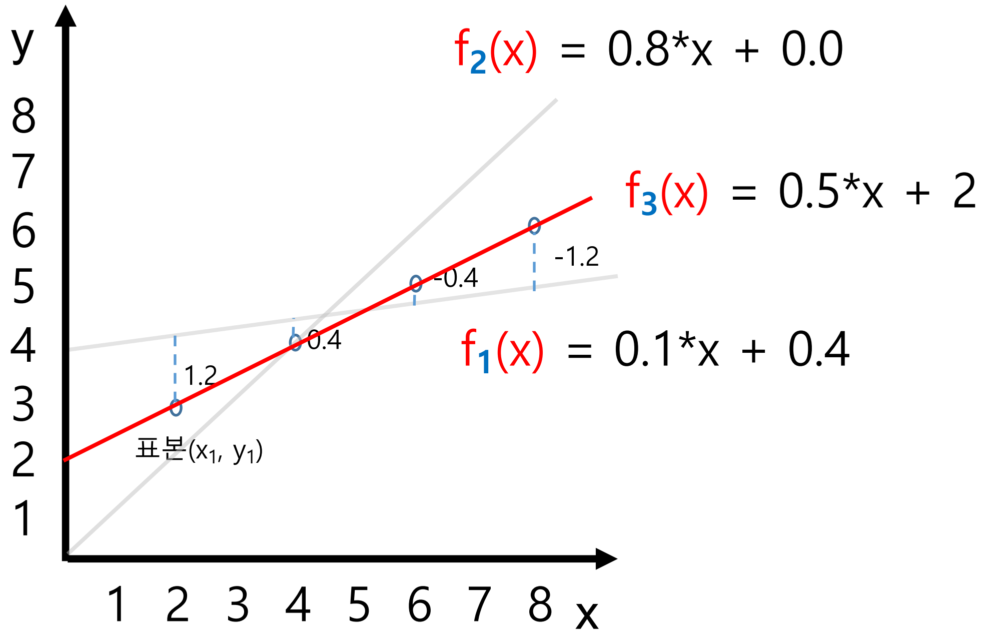


우리가 가정(**H**ypothesis)한 w,b에
- 오차가 많다 = 손실이 많다
- Cost function = Loss function

$$cost(W,b) = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2$$

❑ 최적의 선형회귀란?

- 관측된 표본 데이터를 가장 잘 설명할 수 있는 선을 하나 찾는다. = 최적화된 w, b조합을 찾는다. = Cost(w, b)가 최저 값을 보이는 w, b 값을 찾는다.



Cost(w,b)를 최소화하는 w, b 값을 찾는다.

$$cost(W,b) = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2$$

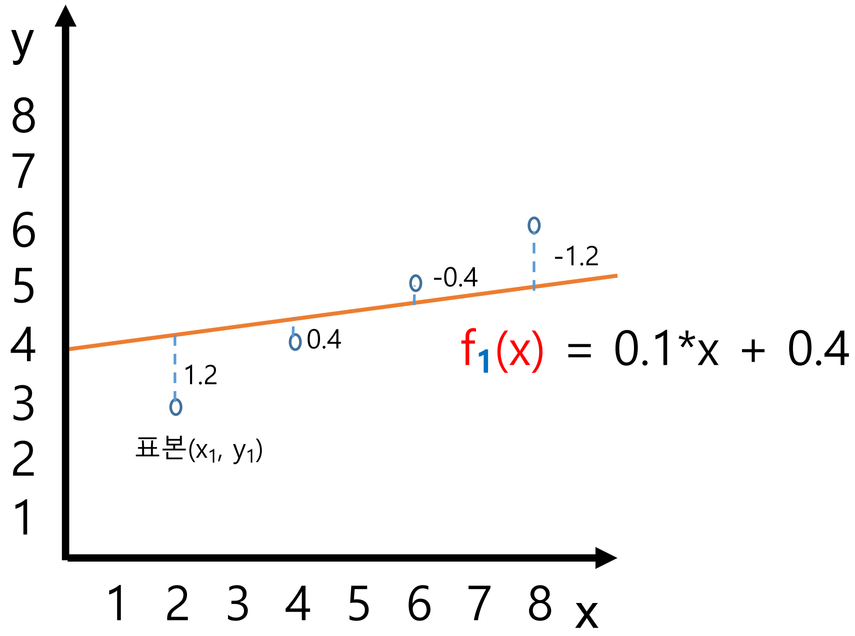
$$\underset{W,b}{\text{minimize}}\ cost(W,b)$$

$$cost(W,b) = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2$$

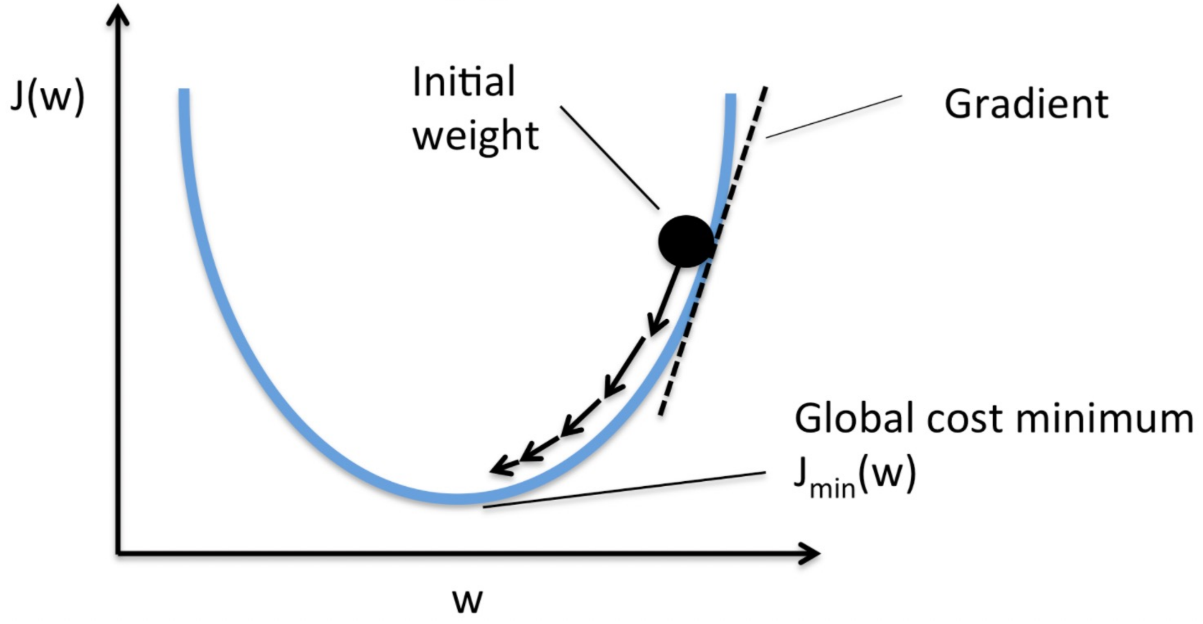
$$\underset{W,b}{\text{minimize}}\; cost(W,b)$$

❑ Cost(w, b)가 최저 값을 보이는 w, b를 어떻게 찾을까?

- w값의 변화에 의해 Cost(w)값이 어떻게 달라지는지 살펴보자
- cost값이 작아지는 **방향**으로 조금씩 w값을 업데이트 하자!!



$$cost(W) = \frac{1}{m}\sum_{i=1}^{m}(Wx^{(i)} - y^{(i)})^2$$



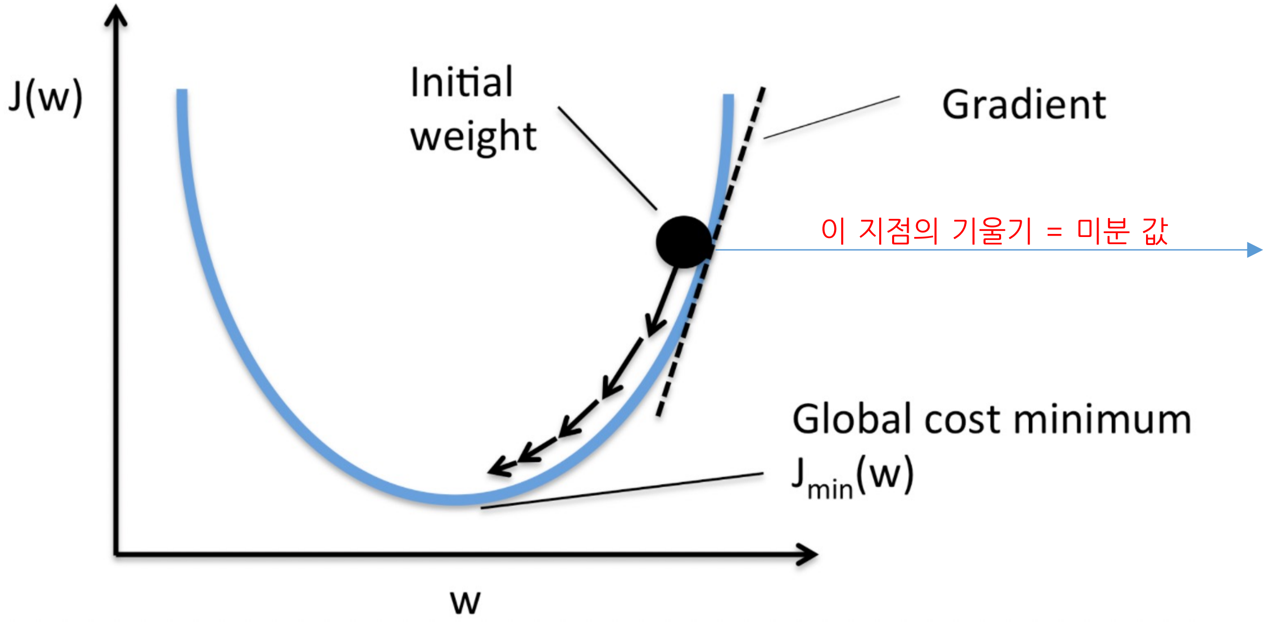
이미지출처: https://sebastianraschka.com/faq/docs/closed-form-vs-gd.html

❑ Cost 값이 작아지는 방향은 어떻게 찾을까?
- 바로 미분! = 기울기

$$\underset{W,b}{\text{minimize}}\, cost(W, b)$$

$$cost(W) = \frac{1}{m}\sum_{i=1}^{m}(Wx^{(i)} - y^{(i)})^2$$

$$\frac{1}{2m}\sum_{i=1}^{m}(Wx^{(i)} - y^{(i)})^2$$



이미지출처: https://sebastianraschka.com/faq/docs/closed-form-vs-gd.html

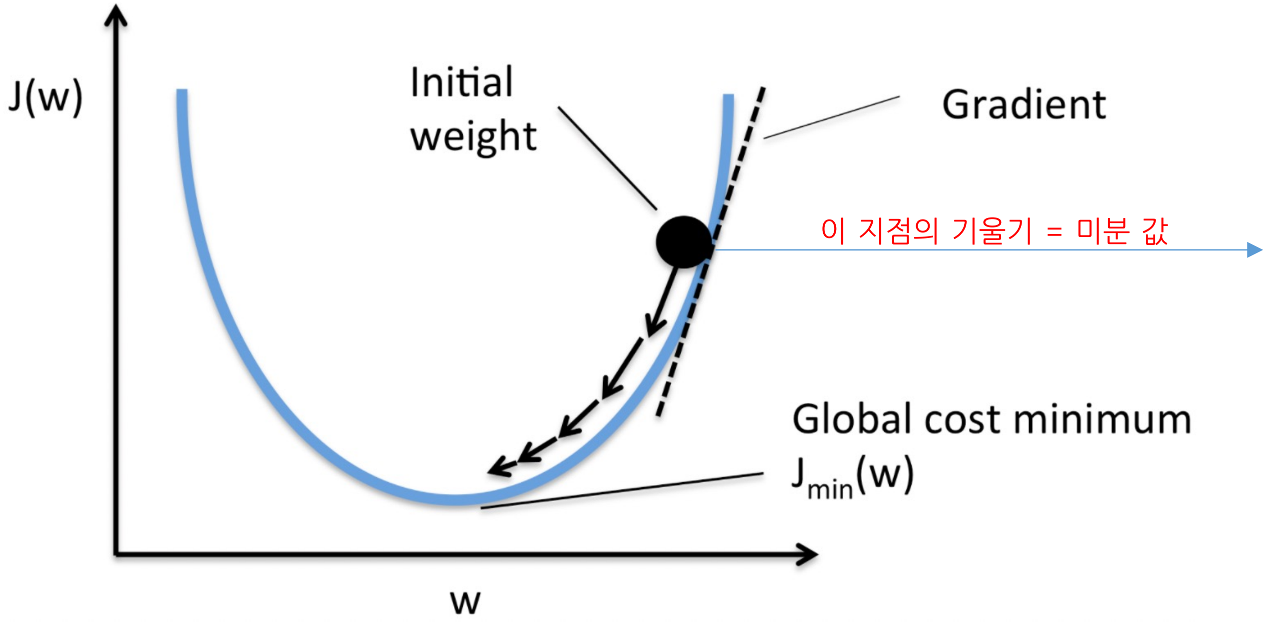
02. Linear Regression (선형회귀 오류 피드백)

❑ Cost 값이 작아지는 방향은 어떻게 찾을까?
- 바로 미분! = 기울기

$$\underset{W,b}{\text{minimize}}\, cost(W, b)$$

$$cost(W) = \frac{1}{m}\sum_{i=1}^{m}(Wx^{(i)} - y^{(i)})^2$$

$$\frac{1}{2m}\sum_{i=1}^{m}(Wx^{(i)} - y^{(i)})^2$$



$$\frac{\partial}{\partial W}cost(W) = \frac{\partial}{\partial W}\frac{1}{2m}\sum_{i=1}^{m}(Wx^{(i)} - y^{(i)})^2$$

$$\frac{1}{2m}\sum_{i=1}^{m}2(Wx^{(i)} - y^{(i)})x^{(i)}$$

$$\frac{1}{m}\sum_{i=1}^{m}(Wx^{(i)} - y^{(i)})x^{(i)}$$

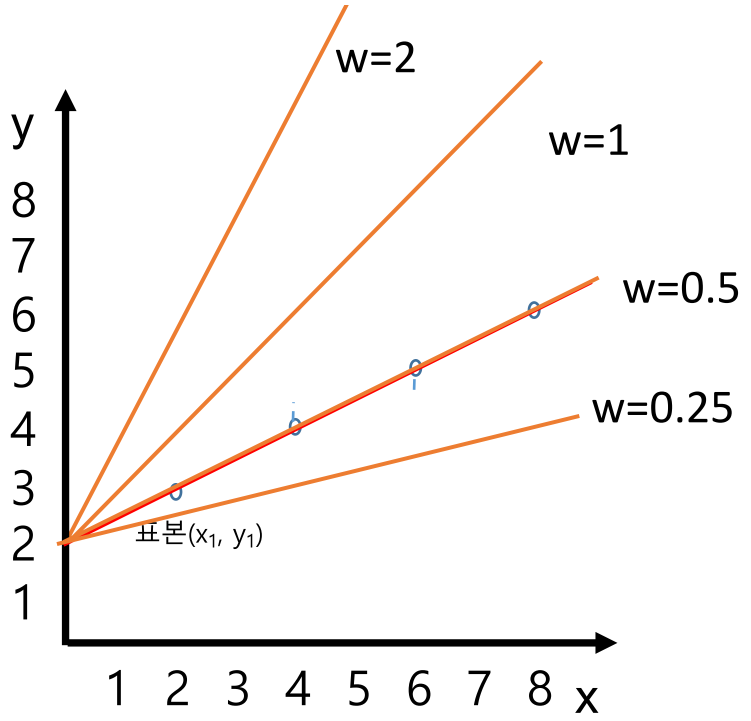
이미지출처: https://sebastianraschka.com/faq/docs/closed-form-vs-gd.html

❑ Gradient Descent Algorithm

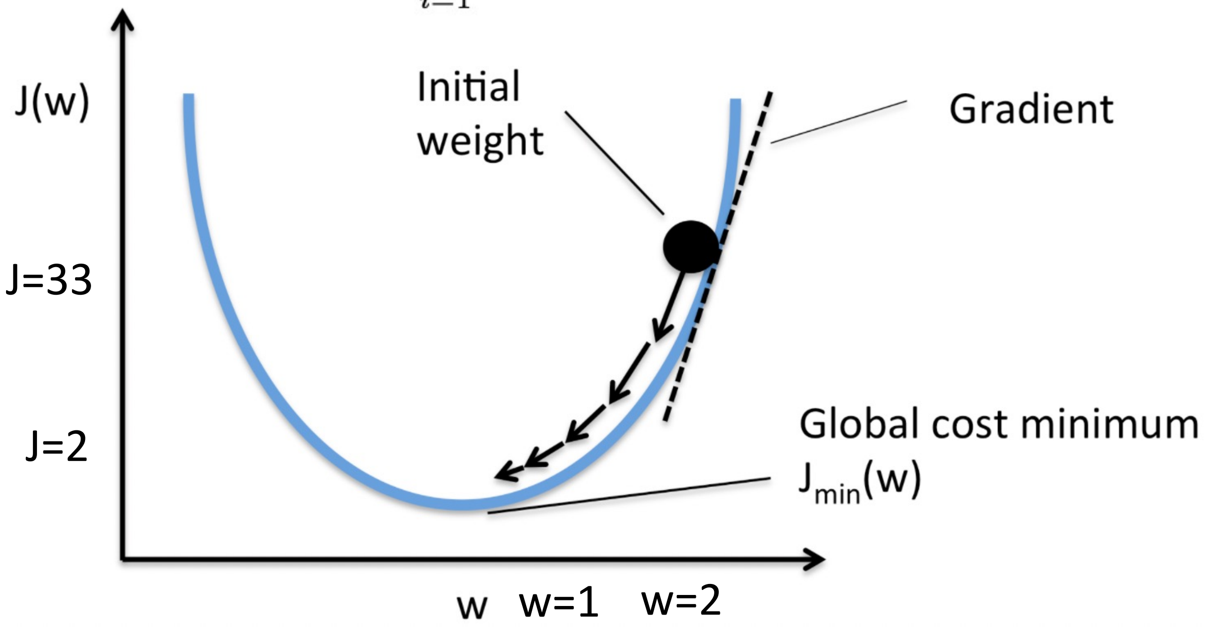
$$\underset{W,b}{\text{minimize}}\; cost(W,b)$$

- 미분을 이용해 cost값이 작아지는 방향으로 조금씩 w값을 업데이트 하자!!

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



$$cost(W) = \frac{1}{m}\sum_{i=1}^{m}(Wx^{(i)} - y^{(i)})^2$$



이미지출처: https://sebastianraschka.com/faq/docs/closed-form-vs-gd.html

02. Linear Regression (선형회귀 오류 피드백)

$$cost(W,b) = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2$$

$$\underset{W,b}{\text{minimize}}\ cost(W,b)$$

❑ Gradient Descent Algorithm

- 미분을 이용해 cost값이 작아지는 방향으로 조금씩 w값을 업데이트 하자!!

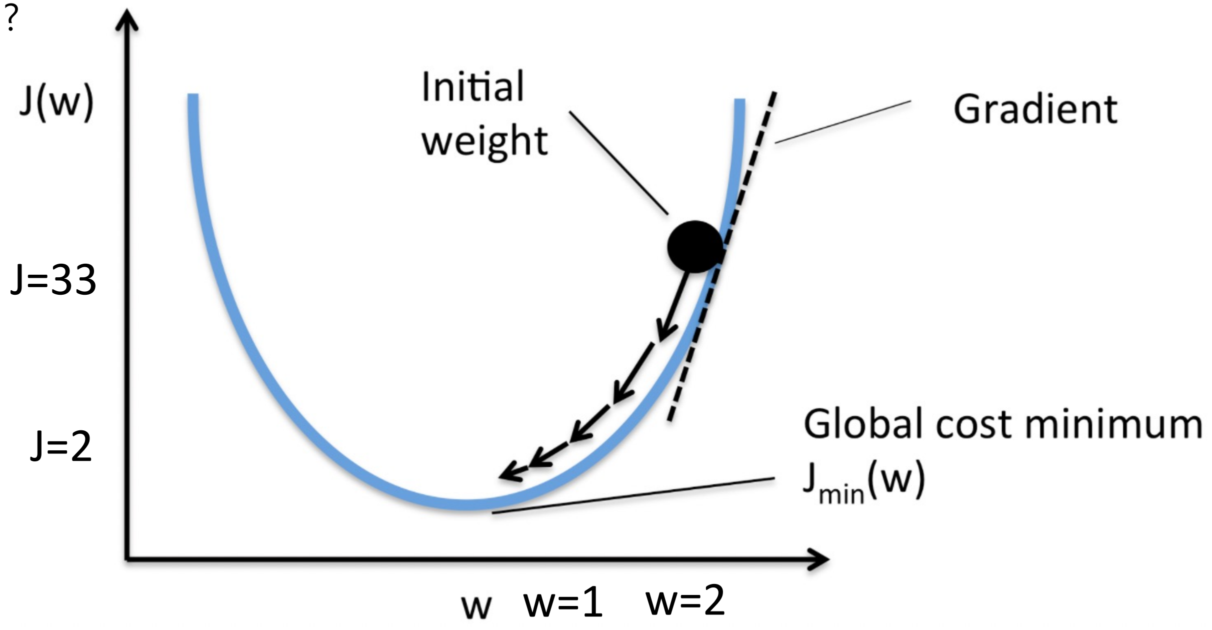
Gradient = 기울기 (미분 값):
w가 cost에 미치는 영향이 얼마나 크니?

현시점W 값

업데이트한
W 값

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

Learning rate =
한번에 얼마만큼 업데이트 할래?



이미지출처: https://sebastianraschka.com/faq/docs/closed-form-vs-gd.html

❑ Gradient Descent Algorithm

- cost값이 작아지는 방향으로 조금씩 **w, b**값을 업데이트 하자!!

Gradient = (미분 값)
w가 cost에 미치는 영향이 얼마나 크니?

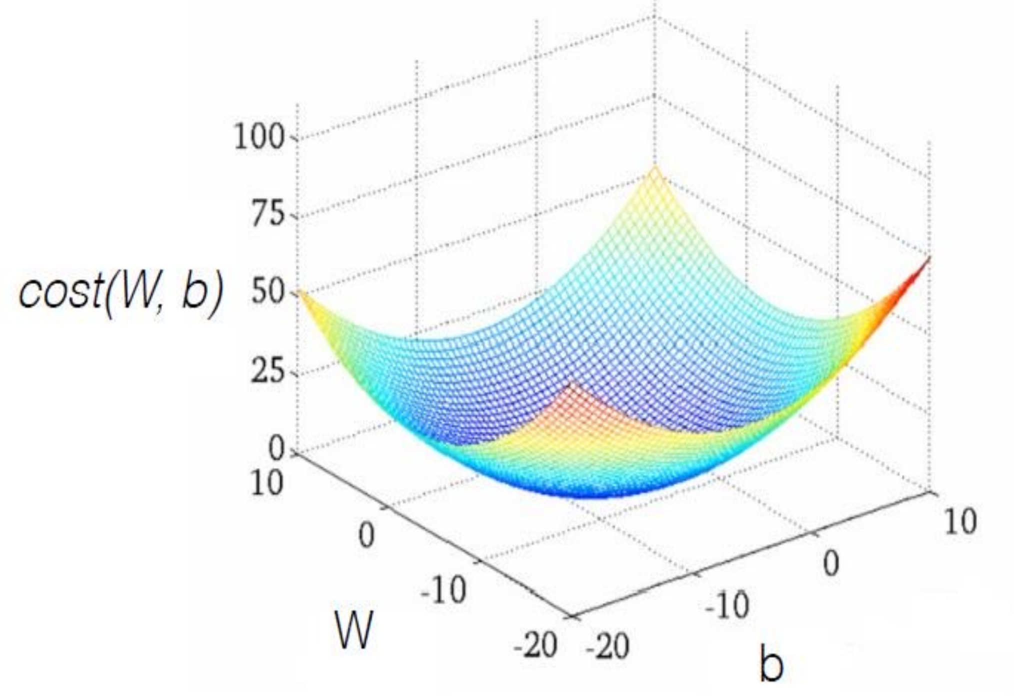
현시점W 값

업데이트한
W 값

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

Learning rate =
한번에 얼마만큼 업데이트 할래?

$$cost(W,b) = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2$$



이미지출처: holehouse.org/mlclass/

02. Linear Regression (선형회귀 오류 피드백)

❑ Gradient Descent Algorithm

- cost값이 작아지는 방향으로 조금씩 **w, b**값을 업데이트 하자!!

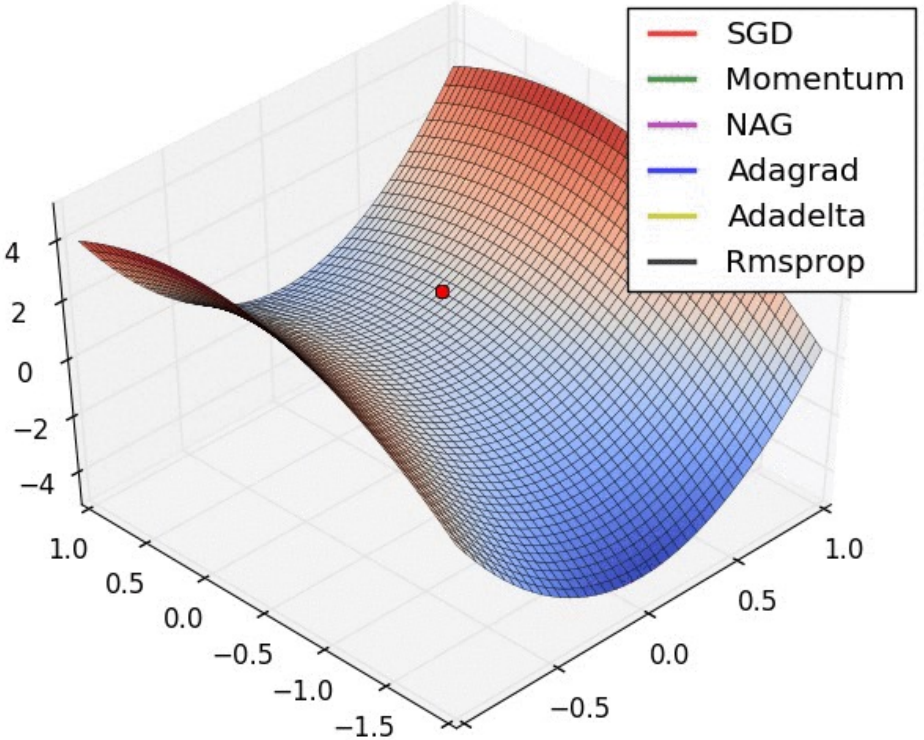
Gradient = (미분 값)
w가 cost에 미치는 영향이 얼마나 크니?

현시점W 값

업데이트한
W 값

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

Learning rate =
한번에 얼마만큼 업데이트 할래?



이미지출처: https://imgur.com/a/Hqolp

02. Linear Regression (선형회귀 정리)

❑ Gradient Descent Algorithm 정리

Gradient = (미분 값)
w가 cost에 미치는 영향이 얼마나 크니?

현시점W 값

업데이트한
W 값

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

Learning rate =
한번에 얼마만큼 업데이트 할래?

1. 가설에 따라 W(weight)을 설정하고 X를 넣어 예측 결과물을 얻는다.	**Prediction**
2. 가설에 따른 결과와 실제 값과의 차이를 계산한다.	**Cost Function**
3. 가설과 실제 값의 차이(에러)가 어떤 값(W)에서 비롯되었는가?	**Differentiation (미분)**
4. 역으로 내려가면서 추정하여 W의 가중치를 계산한다.	**Back Propagation**
5. 추정된 가중치*learning rate 만큼 W 값을 변화시킨다.	**Weight Update**

감사합니다.