

선수 지식 - 통계

최소 제공법과 추세선

최소 제공법과 추세선 | 딥러닝의 기초가 되는 확률 개념 알아보기

강사 나동빈

선수 지식 - 통계

최소 제공법과 추세선

선형 회귀(Linear Regression)

- 장사꾼의 매출 문제: 어느 날, 한 장사꾼의 노동 시간과 매출액이 다음과 같았다.

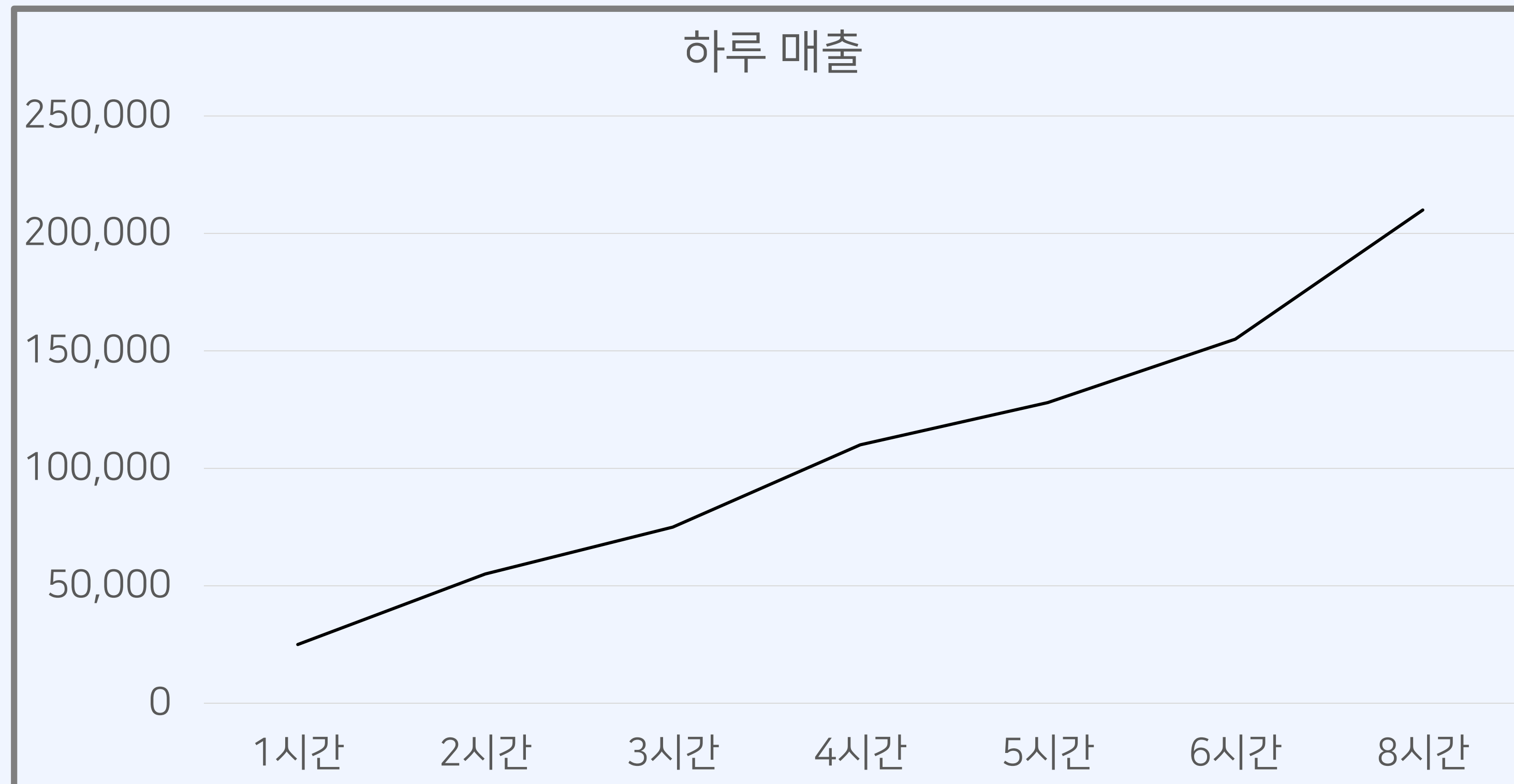
Q. 7시간을 일 했을 때 매출은 얼마일까?

Q. 9시간을 일 했을 때 매출은 얼마일까?

| 하루 노동 시간 | 매출 |
|----------|---------|
| 1 | 25,000 |
| 2 | 55,000 |
| 3 | 75,000 |
| 4 | 110,000 |
| 5 | 128,000 |
| 6 | 155,000 |
| 8 | 210,000 |

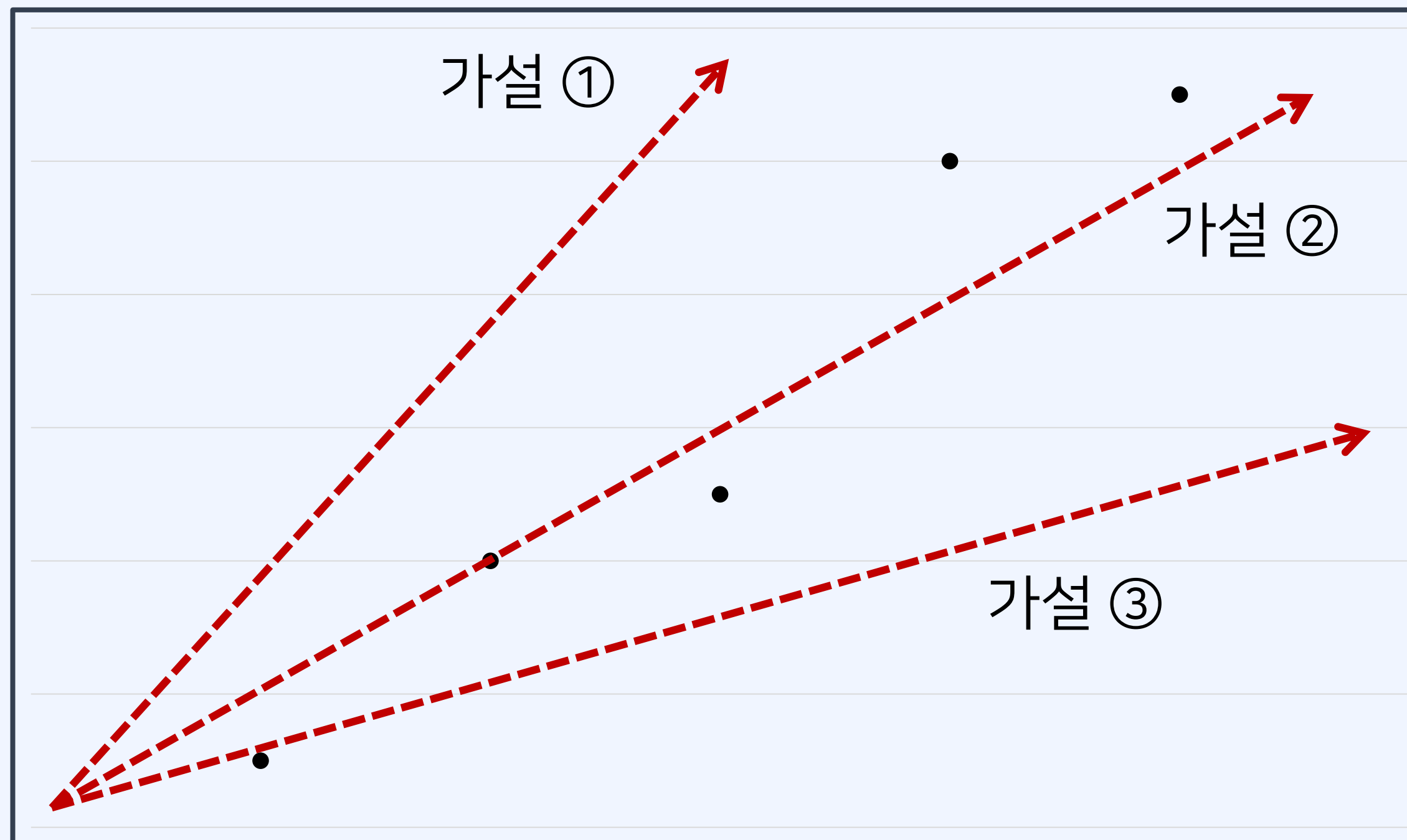
선형 회귀(Linear Regression)

- 매출 데이터를 그래프로 표현해 보자.
- 입력(노동 시간)과 출력(매출)이 선형 함수(일차 함수) 형태를 가진다.



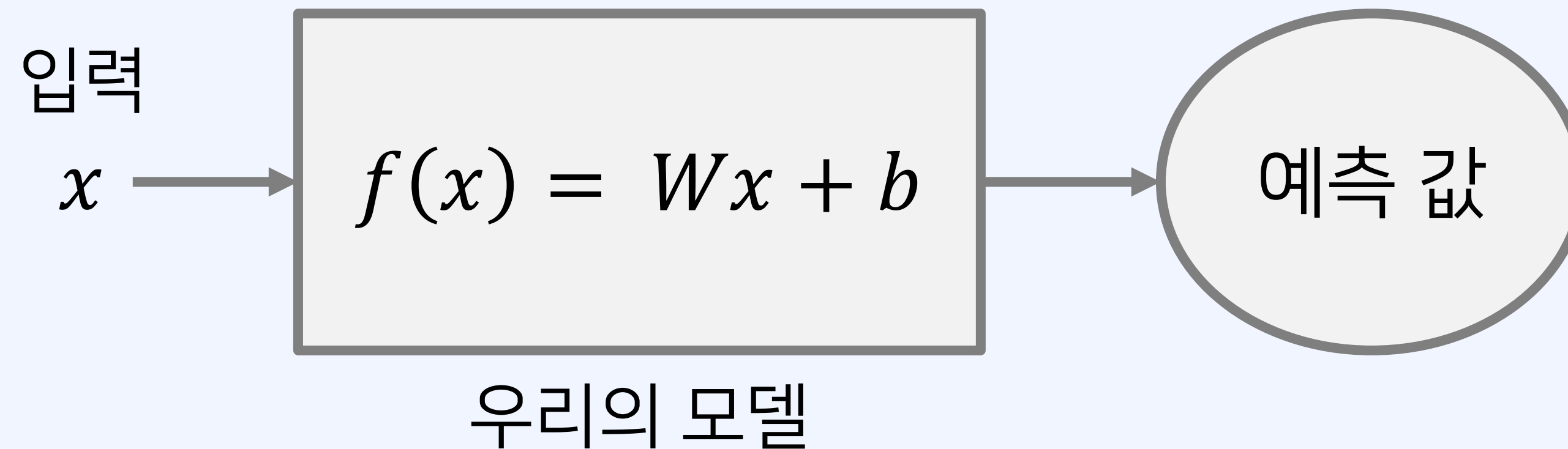
선형 회귀(Linear Regression)

- 선형 회귀: 주어진 데이터를 학습하여 가장 합리적인 선형 함수를 찾아내는 문제다.
- 학습 데이터가 3개 이상일 때 의미가 있다.
- 아래 함수 중에서 데이터를 가장 잘 나타내는 선형 함수는 무엇일까?



선형 회귀(Linear Regression)

- 선형 함수 모델을 가정하므로, 다음과 같이 모델을 정의할 수 있다.
- 이제 우리의 모델(W 와 b)을 수정해 나가면서 가장 합리적인 식을 찾아낼 수 있다.
- 가설 함수: $f(x) = Wx + b$



선형 회귀(Linear Regression)

- 선형 회귀에서의 학습이란?
- 주어진 데이터를 이용해 선형 함수를 수정해 나가는 것이다.
- 학습을 거쳐 가장 합리적인 파라미터(W 와 b)를 도출할 수 있다.

선형 회귀(Linear Regression)

- 현재 우리의 모델이 있을 때, 모델이 얼마나 잘못되었는지 평가하는 방법이 필요하다.
- **비용(cost)**: 우리의 모델이 받은 답이 실제 정답과 얼마나 다른지 수치화한 것
- 우리의 모델이 정확하지 않다면, 높은 비용이 발생한다.
- 비용을 줄이는 방향으로 우리의 모델을 바꾸어 나갈 수 있다.
- 비용은 다른 말로 **손실(loss)**이라고도 한다.

최소제곱법(Least Square Method)

- 평균 제곱 오차(mean squared error)를 이용할 수 있다.
- 모든 데이터에 대한 (실제 값 - 예상 값)²의 합으로 비용을 계산한다.
- 따라서 다음 식을 **최소화**하는 파라미터 (W, b)를 찾는 것이 목표다.
→ 이를 최소제곱법이라고 한다.
- m : 데이터의 개수

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

최소제곱법(Least Square Method)

비용(cost) 구하기 예제

- $W = 1, b = 2$ 일 때의 비용을 계산해 보자.
- $y = x + 2$

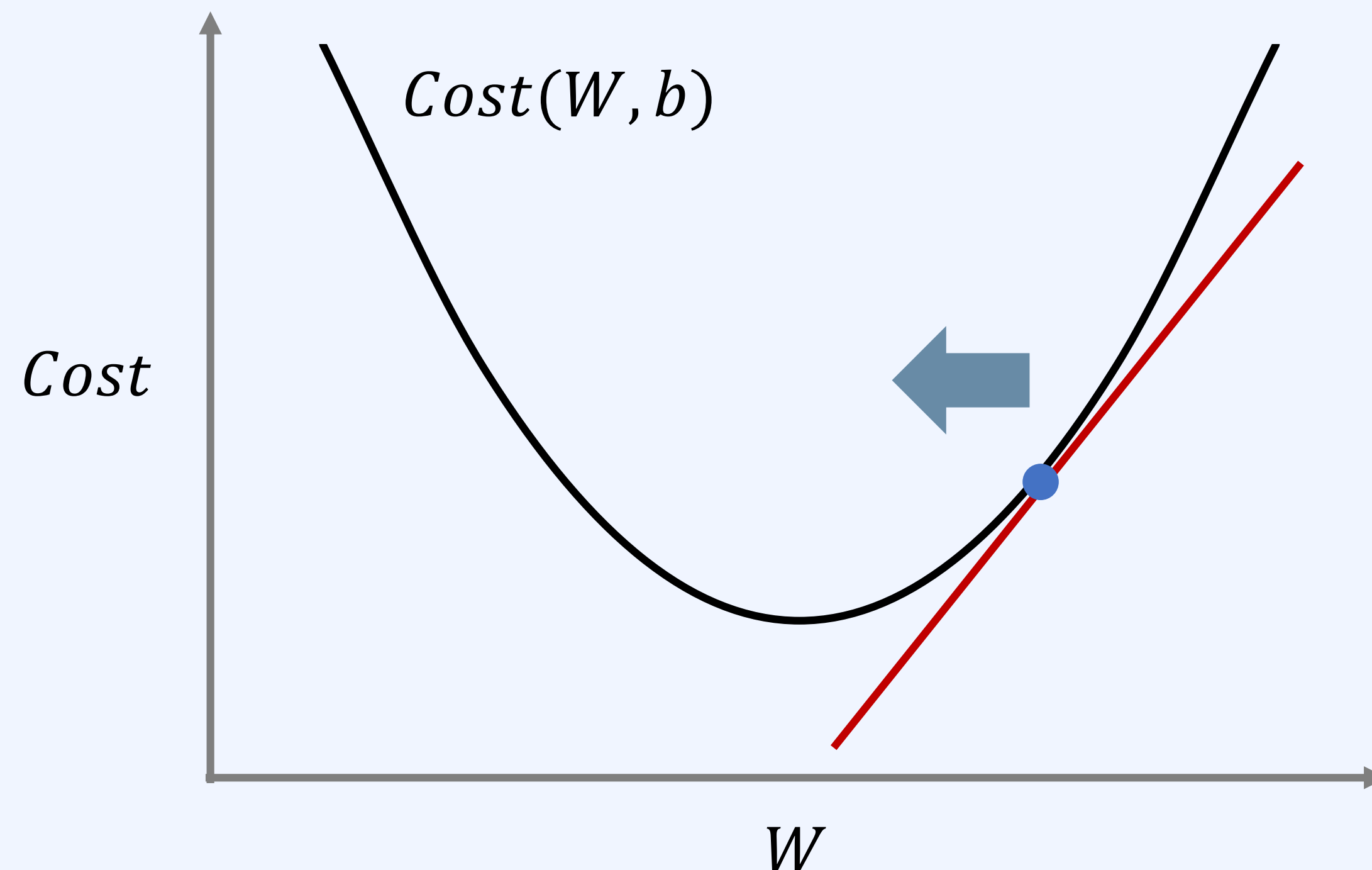
| X | Y |
|---|---|
| 1 | 2 |
| 2 | 4 |
| 3 | 6 |

$$\begin{aligned}
 cost(W, b) &= \frac{1}{m} \sum_{i=1}^m (H(x_i) - y_i)^2 \\
 &= \frac{(3 - 2)^2 + (4 - 4)^2 + (5 - 6)^2}{3} \\
 &= 2/3
 \end{aligned}$$

➡ 비용을 더 줄려면, W 와 b 를 어떻게 바꾸어야 할까?

최소 제곱 값을 얻는 방법 - 경사 하강

- 최소 제곱을 얻는 방법 중 하나는 바로 “경사 하강”을 이용하는 것이다.
- 경사 하강(gradient descent): 기울기를 이용하여 비용을 줄이는 방법이다.
- 미분을 이용하면 특정 값에서의 기울기를 구할 수 있다.



“현재 기울기가 양수(+)구나?
가중치를 음수(-) 방향으로 이동시키자!”

- 간단히 $H(x) = Wx$ 라고 생각해 보자.
- 비용 함수 MSE는 다음과 같다.

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

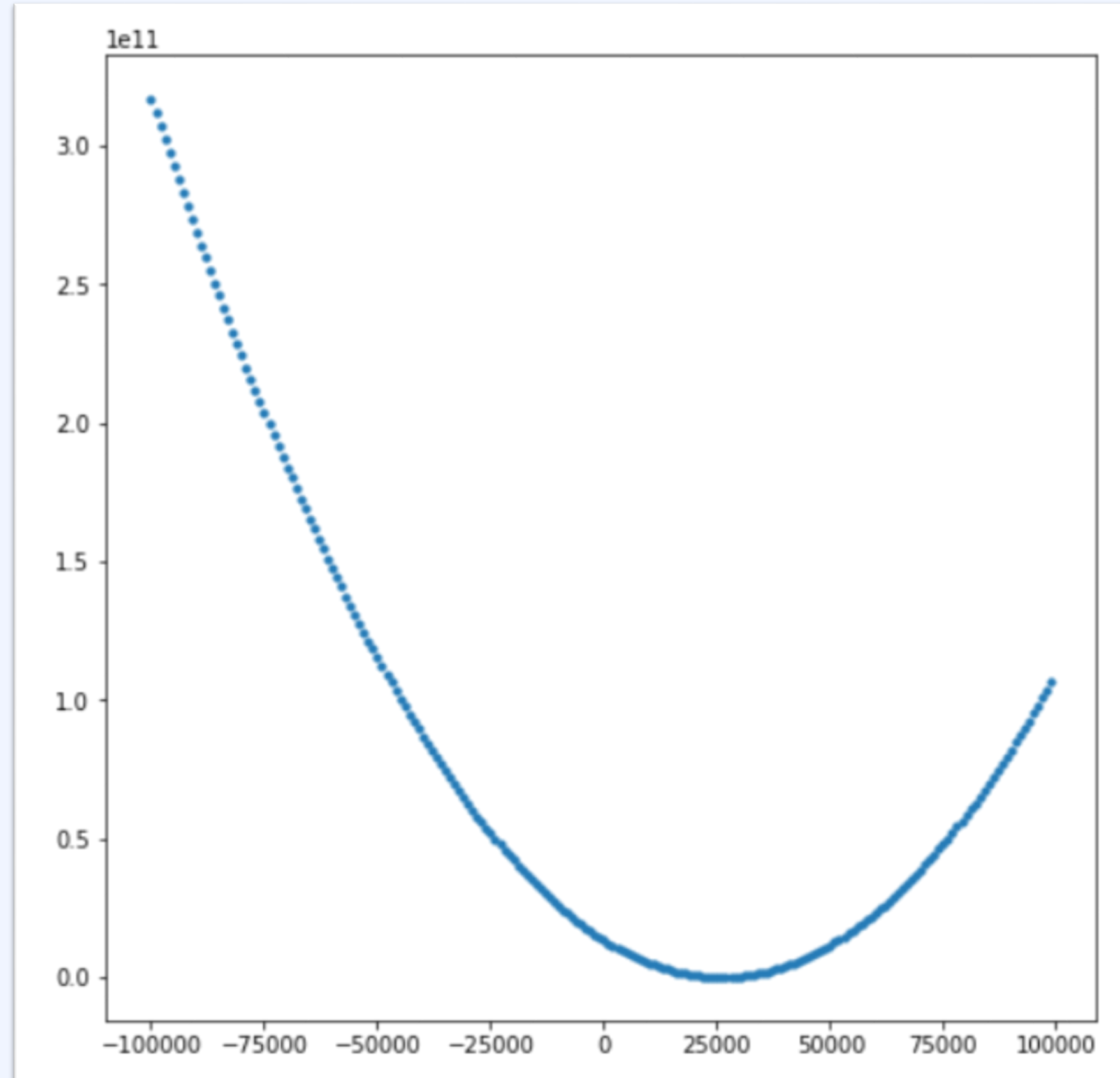
- 그렇다면 $(Wx_i - y_i)^2$ 의 합을 어떻게 하면 가장 작게 할 수 있을까?

선수 지식 - 통계
최소 제곱법과 추세선

최소 제곱 값을 얻는 방법 - 경사 하강

선수 지식
통계
최소 제곱법과
추세선

- MSE 비용 함수는 볼록 함수로, 경사를 타고 내려가면 비용이 감소한다.



최소제곱법으로 추세선 찾기

- 다시 말해, 비용을 줄이기 위한 방법은 다음과 같다.
→ 비용 함수를 가중치(W)로 미분해 기울기를 구한 뒤에, 기울기의 반대 방향으로 W 를 업데이트
- 어느 정도의 크기로 이동해야 할까?
- 적절한 크기로 이동할 수 있도록, 학습률(learning rate)을 곱하여 이동한다.
- 예를 들어 학습률이 0.01이고, 기울기가 7이라면?
- $W \leftarrow W - 7 * 0.01$ 로 업데이트한다.
- 학습 과정을 반복하면, 가중치는 올바른 지점으로 수렴할 수 있다.