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

03. Multi-Variable Linear Regression (개념)

❑ 기계학습 예제를 위해 가장 간단한 형태의 데이터를 고려해보자

- 친한 친구 4명의 수학, 영어 공부시간 대비 점수가 아래와 같다. 내가 4, 7시간 공부하면 몇 점 받을 수 있을까?
- 어떤 과목 공부가 수능 점수에 영향을 많이 미치는가?

이름	수학 공부시간	영어 공부시간	과학 공부시간	수능 점수
야쓰오	2	3	1	3
오공	4	8	9	7
블리츠	6	2	7	5
페이커	8	4	6	8

03. Multi-Variable Linear Regression (개념)

❑ 기계학습을 위해 모델링 예제로 바꿔보자

- 어떻게 모델링을 해야할까?

$X = \{x_1, x_2, x_3\}$ where

x_1 = 수학 공부시간 {2,4,6,8}

x_2 = 영어 공부시간 {3,8,2,4}

x_3 = 과학 공부시간 {1,9,7,6}

입력 X값에 **어떤 변환**을 하니까 y가 나온다

$f(x) = ??$

Y = 점수 {3,6,5,8}

03. Multi-Variable Linear Regression (Single vs Multi 비교)

❑ 모델링 방법1: **선형회귀**

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

X = 수학 공부시간 {2,4,6,8}

입력 x값에 **선형 변환**을 하니까 y가 나온다

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

Y = 점수 {3,4,5,6}

$X = \{x_1, x_2, x_3\}$ where

x_1 = 수학 공부시간 {2,4,6,8}

x_2 = 영어 공부시간 {3,8,2,4}

x_3 = 과학 공부시간 {1,9,7,6}

$f(x) = w_1*x_1 + w_2*x_2 + w_3*x_3 + b$

Y = 점수 {3,6,5,8}

03. Multi-Variable Linear Regression (Single vs Multi 비교)

❑ 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.

	Single Variable	Multi Variable
Hypothesis	$H(x) = Wx + b$	$H(x_1, x_2, x_3) = w_1x_1 + w_2x_2 + w_3x_3 + b$
Cost function	$cost(W,b) = \frac{1}{m}\sum_{i=1}^{m}(H(x^{(i)}) - y^{(i)})^2$	$cost(W,b) = \frac{1}{m}\sum_{I=1}^{m}(H(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}) - y^{(i)})^2$
Goal	$\underset{W,b}{\text{minimize}}\, cost(W,b)$	$\underset{W,b}{\text{minimize}}\, cost(W,b)$ *Where W = {w_1,w_2,w_3}*

03. Multi-Variable Linear Regression (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 **각 w값**을 업데이트 하자!!

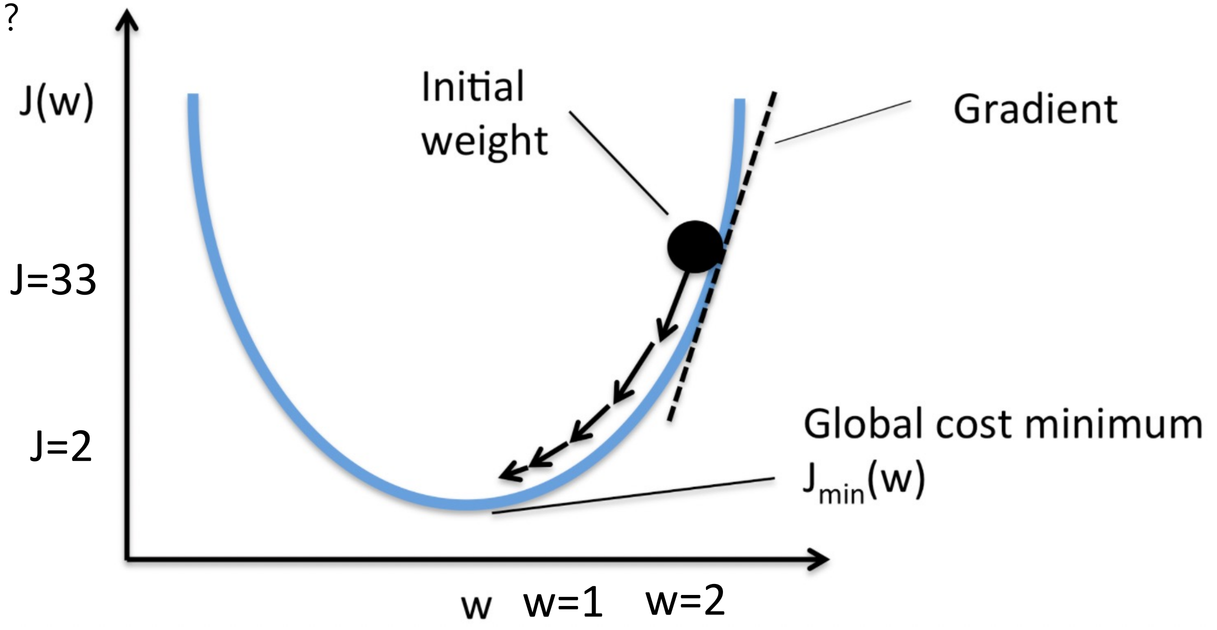
Gradient = 기울기 (미분 값):
w가 cost에 미치는 영향이 얼마나 크니?

현시점W 값

업데이트한
W 값

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

Learning rate =
한번에 얼마만큼 업데이트 할래?



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

❑ Multi-Variable 의 연산

- Variable의 개수와 상관없이 결국 행렬 연산으로 쉽게 계산이 가능한 구조가 나온다.

$$H(x_1, x_2, x_3) = w_1x_1 + w_2x_2 + w_3x_3 + b$$

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} x_1w_1 + x_2w_2 + x_3w_3 \end{pmatrix}$$

$$H(X) = XW$$

❑ Multi-Variable 의 연산

- Variable의 개수와 상관없이 결국 행렬 연산으로 쉽게 계산이 가능한 구조가 나온다.

$$H(x_1, x_2, x_3) = w_1x_1 + w_2x_2 + w_3x_3 + b$$

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} x_1w_1 + x_2w_2 + x_3w_3 \end{pmatrix}$$

$$H(X) = XW$$

❑ Multi-Variable 의 연산

- Variable의 개수와 상관없이 결국 행렬 연산으로 쉽게 계산이 가능한 구조가 나온다.

Sample이 **한** 개일 때

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} x_1 w_1 + x_2 w_2 + x_3 w_3 \end{pmatrix}$$

$$H(X) = XW$$

Sample이 **다섯** 개일 때

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \\ x_{51} & x_{52} & x_{53} \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} x_{11} w_1 + x_{12} w_2 + x_{13} w_3 \\ x_{21} w_1 + x_{22} w_2 + x_{23} w_3 \\ x_{31} w_1 + x_{32} w_2 + x_{33} w_3 \\ x_{41} w_1 + x_{42} w_2 + x_{43} w_3 \\ x_{51} w_1 + x_{52} w_2 + x_{53} w_3 \end{pmatrix}$$

$$H(X) = XW$$

❑ Multi-Variable 의 연산

- Variable의 개수와 상관없이 결국 행렬 연산으로 쉽게 계산이 가능한 구조가 나온다.

Sample이 **한** 개일 때

$$\begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} x_1 w_1 + x_2 w_2 + x_3 w_3 \end{pmatrix}$$

$$H(X) = XW$$

이름	수학	영어	과학	점수
야쓰오	2	3	1	3

Sample이 **다섯** 개일 때

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \\ x_{51} & x_{52} & x_{53} \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} x_{11}w_1 + x_{12}w_2 + x_{13}w_3 \\ x_{21}w_1 + x_{22}w_2 + x_{23}w_3 \\ x_{31}w_1 + x_{32}w_2 + x_{33}w_3 \\ x_{41}w_1 + x_{42}w_2 + x_{43}w_3 \\ x_{51}w_1 + x_{52}w_2 + x_{53}w_3 \end{pmatrix}$$

$$H(X) = XW$$

이름	수학	영어	과학	점수
야쓰오	2	3	1	3
오공	4	8	9	7
블리츠	6	2	7	5
페이커	8	4	6	8

03. Multi-Variable Linear Regression (Single-sample vs Multi-sample 비교)

❑ 행렬 연산이 왜 중요한가요?

- **병렬 연산 가능!**

Sample이 **다섯** 개일 때

$$\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \\ x_{51} & x_{52} & x_{53} \end{pmatrix} \cdot \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = \begin{pmatrix} x_{11}w_1 + x_{12}w_2 + x_{13}w_3 \\ x_{21}w_1 + x_{22}w_2 + x_{23}w_3 \\ x_{31}w_1 + x_{32}w_2 + x_{33}w_3 \\ x_{41}w_1 + x_{42}w_2 + x_{43}w_3 \\ x_{51}w_1 + x_{52}w_2 + x_{53}w_3 \end{pmatrix}$$

$$H(X) = XW$$

```
>>> a = [[1, 0], [0, 1]]
>>> b = [[4, 1], [2, 2]]
>>> np.dot(a, b)
array([[4, 1],
       [2, 2]])
```

감사합니다.