Lecture 4

Multi-variable linear regression

Sung Kim <hunkim+mr@gmail.com>

Recap

- Hypothesis

- Cost function

- Gradient descent algorithm

Recap

- Hypothesis

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

- Cost function

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

- Gradient descent algorithm

Predicting exam score: regression using one input (x)

one-variable
one-feature

x (hours)	y (score)
10	90
9	80
3	50
2	60
11	40

Predicting exam score: regression using two inputs (x1, x2)

multi-variable/feature

x1 (hours)	x2 (attendance)	y (score)
10	5	90
9	5	80
3	2	50
2	4	60
11	1	40

Hypothesis

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

Hypothesis

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

$$H(x_1, x_2) = w_1 x_1 + w_2 x_2 + b$$

Cost function

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

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

Multi-variable

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

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

Matrix

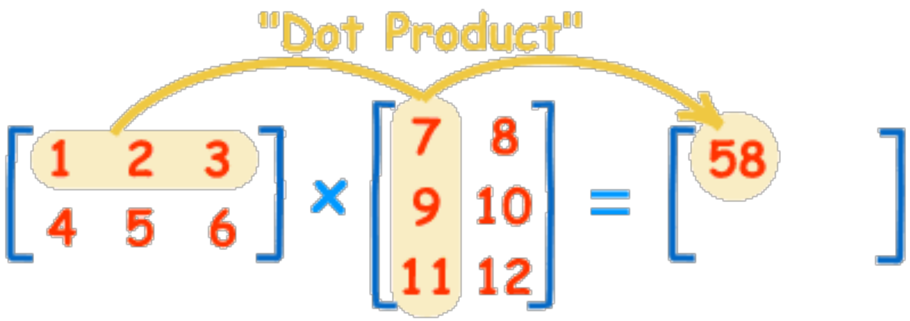
$$w_1x_1 + w_2x_2 + w_3x_3 + ... + w_nx_n$$

Matrix

$$w_1x_1 + w_2x_2 + w_3x_3 + ... + w_nx_n$$

$$\begin{bmatrix} w1 & w2 & w3 \end{bmatrix} \times \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} w1 \times x1 + w2 \times x2 + w3 \times x3 \end{bmatrix}$$

Matrix multiplication



https://www.mathsisfun.com/algebra/matrix-multiplying.html

Hypothesis

$$\begin{bmatrix} w1 & w2 & w3 \end{bmatrix} \times \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} w1 \times x1 + w2 \times x2 + w3 \times x3 \end{bmatrix}$$

$$H(X) = WX + b$$

Hypothesis without *b*

$$\begin{bmatrix} b & w1 & w2 & w3 \end{bmatrix} \times \begin{bmatrix} 1 \\ x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} b \times 1 + w1 \times x1 + w2 \times x2 + w3 \times x3 \end{bmatrix}$$

$$H(X) = WX$$

W vs X

$$\begin{bmatrix} b & w1 & w2 & w3 \end{bmatrix} \times \begin{bmatrix} 1 \\ x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} b \times 1 + w1 \times x1 + w2 \times x2 + w3 \times x3 \end{bmatrix}$$

$$H(X) = WX$$

Transpose

$$\begin{bmatrix} 6 & 4 & 24 \\ 1 & -9 & 8 \end{bmatrix}^T = \begin{bmatrix} 6 & 1 \\ 4 & -9 \\ 24 & 8 \end{bmatrix}$$

http://www.mathsisfun.com/algebra/matrix-introduction.html

Hypothesis using Transpose

$$\begin{bmatrix} b & w1 & w2 & w3 \end{bmatrix} \times \begin{bmatrix} 1 \\ x1 \\ x2 \\ x3 \end{bmatrix} = \begin{bmatrix} b \times 1 + w1 \times x1 + w2 \times x2 + w3 \times x3 \end{bmatrix}$$

$$H(X) = W^T X$$

Next

Logistic Regression
(Classification)