LECTURE 4

MULTIVARIABLE LOGISTIC REGRESSION

Sung Kim <hunkim+ml@gmail.com>
http://hunkim.github.io/ml

Recap

01. Hypothesis

02. Cost function

03. Gradient descent algorithm

Recap

01. Hypothesis
    - $H(x) = Wx + b$

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

03. Gradient descent algorithm

Predicting Exam Score

Regression using one input (X)
One-variable/feature

X (hours)	Y (score)
10	90
9	80
3	50
2	60
11	40

Predicting Exam Score

Regression using three inputs (X1,X2,X3)
Multi-variable/feature

Test Scores for General Psychology

x_1	x_2	x_3	Y
73	80	75	152
93	88	93	185
89	91	90	180
96	98	100	196
73	66	70	142

Hypothesis

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

Hypothesis

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

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

Cost Function

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

$$cost(W, b) = \frac{1}{m}\sum_{i=1}^{m}\left(H(x_1^{(i)}, x_2^{(i)}, x_3^{(i)}) - y^{(i)}\right)^2$$

Multi-variable

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

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

Matrix

$$w_1x_1 + w_2x_2 + w_3x_3 + \cdots + w_nx_n$$

Matrix Multiplication

"Dot Product"

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \times \begin{bmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{bmatrix} = \begin{bmatrix} 58 & \\ & \end{bmatrix}$$

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

Hypothesis Using Matrix

$$w_1x_1 + w_2x_2 + w_3x_3 + \cdots + w_nx_n$$

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

$$H(X) = XW$$

Many x Instances

Test Scores for General Psychology

X_1	X_2	X_3	Y
73	80	75	152
93	88	93	185
89	91	90	180
96	98	100	196
73	66	70	142

http://college.cengage.com/mathematics/brase/understandable_statistics/7e/students/datasets/mlr/frames/frame.html

Hypothesis Using Matrix

$$H(x_1, x_2, x_3) = x_1w_1 + x_2w_2 + x_3w_3$$

Test Scores for General Psychology

x_1	x_2	x_3	Y
73	80	75	152
93	88	93	185
89	91	90	180
96	98	100	196
73	66	70	142

https://www.symbolab.com/solver/matrix-calculator

Hypothesis Using Matrix

$$H(x_1, x_2, x_3) = x_1w_1 + x_2w_2 + x_3w_3$$

Test Scores for General Psychology

x_1	x_2	x_3	Y
73	80	75	152
93	88	93	185
89	91	90	180
96	98	100	196
73	66	70	142

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

$$H(X) = XW$$

https://www.symbolab.com/solver/matrix-calculator

Hypothesis Using Matrix

$$w_1x_1 + w_2x_2 + w_3x_3 + \cdots + w_nx_n$$

Test Scores for General Psychology

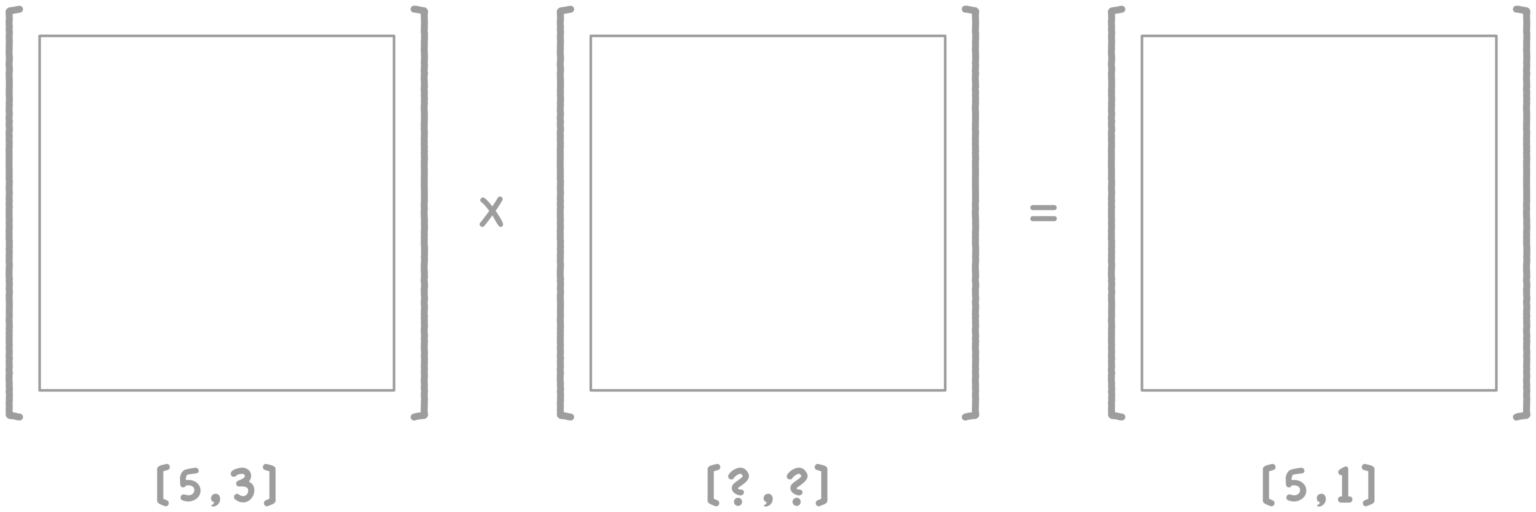
x_1	x_2	x_3	Y
73	80	75	152
93	88	93	185
89	91	90	180
96	98	100	196
73	66	70	142

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

$$H(X) = XW$$

https://www.symbolab.com/solver/matrix-calculator

Hypothesis Using Matrix



$$H(X) = XW$$

Hypothesis Using Matrix

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

[5,3] [3,1] [5,1]

$$H(X) = XW$$

https://www.symbolab.com/solver/matrix-calculator

Hypothesis Using Matrix

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

[n, 3] [3, 1] [n, 1]

$$H(X) = XW$$

https://www.symbolab.com/solver/matrix-calculator

Hypothesis Using Matrix (n output)

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \\ x_{51} & x_{52} & x_{53} \end{bmatrix} \times \begin{bmatrix} \quad \end{bmatrix} = \begin{bmatrix} x_{11}w_{11} + x_{12}w_{21} + x_{13}w_{31} & x_{11}w_{12} + x_{12}w_{22} + x_{13}w_{32} \\ x_{21}w_{11} + x_{22}w_{21} + x_{23}w_{31} & x_{21}w_{12} + x_{22}w_{22} + x_{23}w_{32} \\ x_{31}w_{11} + x_{32}w_{21} + x_{33}w_{31} & x_{31}w_{12} + x_{32}w_{22} + x_{33}w_{32} \\ x_{41}w_{11} + x_{42}w_{21} + x_{43}w_{31} & x_{41}w_{12} + x_{42}w_{22} + x_{43}w_{32} \\ x_{51}w_{11} + x_{52}w_{21} + x_{53}w_{31} & x_{51}w_{12} + x_{52}w_{22} + x_{53}w_{32} \end{bmatrix}$$

[n, 3] [?, ?] [n, 2]

$$H(X) = XW$$

https://www.symbolab.com/solver/matrix-calculator

Hypothesis Using Matrix (n output)

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \\ x_{41} & x_{42} & x_{43} \\ x_{51} & x_{52} & x_{53} \end{bmatrix} \times \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \\ w_{31} & w_{32} \end{bmatrix} = \begin{bmatrix} x_{11}w_{11} + x_{12}w_{21} + x_{13}w_{31} & x_{11}w_{12} + x_{12}w_{22} + x_{13}w_{32} \\ x_{21}w_{11} + x_{22}w_{21} + x_{23}w_{31} & x_{21}w_{12} + x_{22}w_{22} + x_{23}w_{32} \\ x_{31}w_{11} + x_{32}w_{21} + x_{33}w_{31} & x_{31}w_{12} + x_{32}w_{22} + x_{33}w_{32} \\ x_{41}w_{11} + x_{42}w_{21} + x_{43}w_{31} & x_{41}w_{12} + x_{42}w_{22} + x_{43}w_{32} \\ x_{51}w_{11} + x_{52}w_{21} + x_{53}w_{31} & x_{51}w_{12} + x_{52}w_{22} + x_{53}w_{32} \end{bmatrix}$$

[n, 3] [3, 2] [n, 2]

$$H(X) = XW$$

https://www.symbolab.com/solver/matrix-calculator

WX vs XW

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

Lecture
(Theory)

$$H(x)=XW$$

Implementation
(TensorFlow)

Hypothesis without b

$$\begin{bmatrix} b & w_1 & w_2 & w_3 \end{bmatrix} \times \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b \times 1 + w_1 \times x_1 + w_2 \times x_2 + w_3 \times x_3 \end{bmatrix}$$

$$H(X) = WX$$

W vs X

$$\begin{bmatrix} b & w_1 & w_2 & w_3 \end{bmatrix} \times \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b \times 1 + w_1 \times x_1 + w_2 \times x_2 + w_3 \times x_3 \end{bmatrix}$$

$$H(X) = WX$$

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

$$\begin{bmatrix} b & w_1 & w_2 & w_3 \end{bmatrix} \times \begin{bmatrix} 1 \\ x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b \times 1 + w_1 \times x_1 + w_2 \times x_2 + w_3 \times x_3 \end{bmatrix}$$

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

NEXT LECTURE

LOGISTIC REGRESSION (CLASSIFICATION)