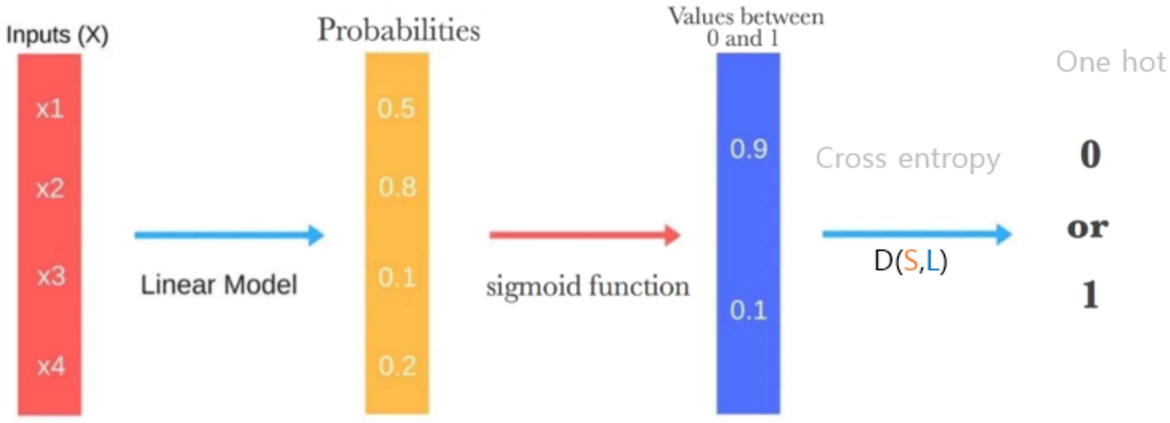
# 4. Softmax Classifier

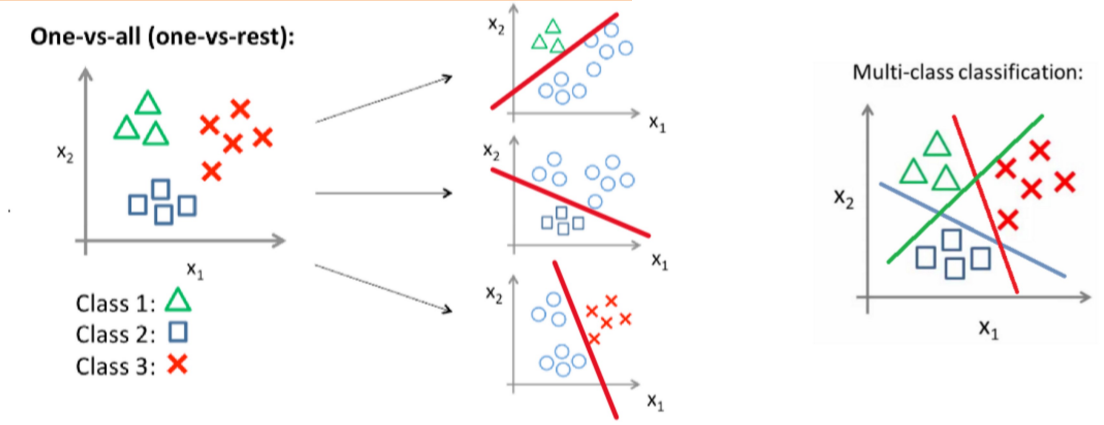
2019년 3월 29일 금요일 오후 4:43

## 1. ***Definition***

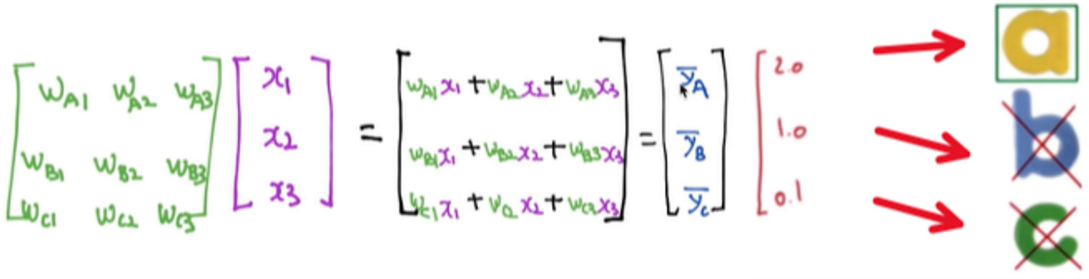
- N개의 결과값을 추측할 때



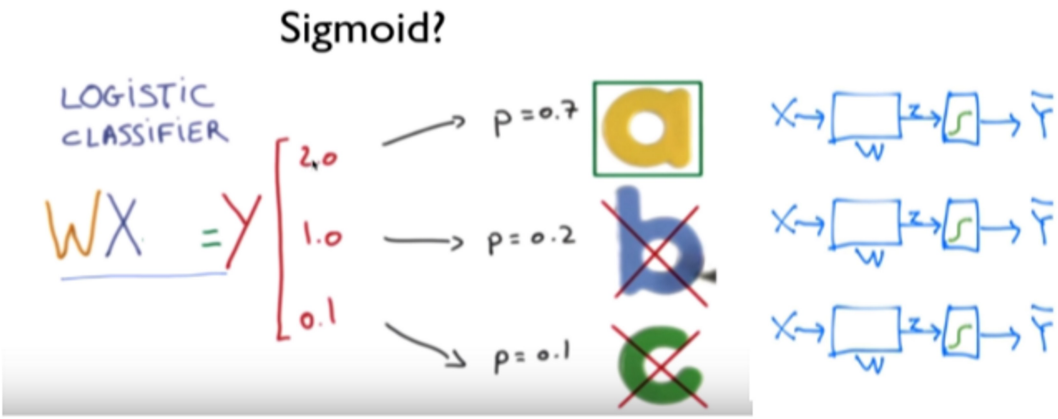

- ***Multinomial classification : 데이터 -> 분류***



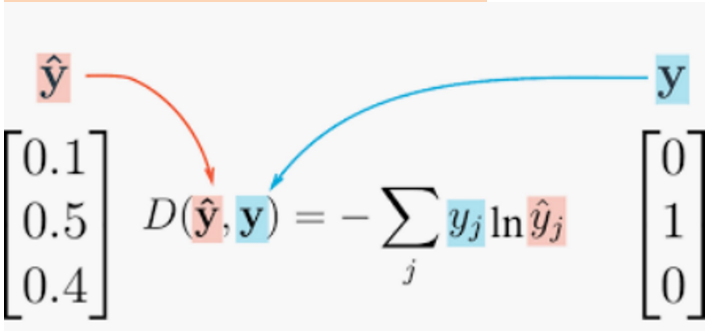

- ***Linear model : 데이터 -> 값***



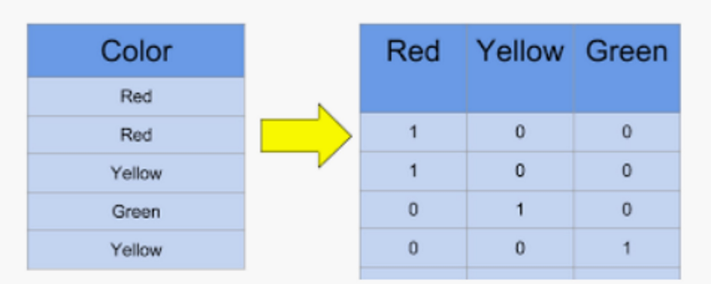
- ***Sigmoid : 값 -> 확률***
  - 0 < p < 1 : 확률
  - P값의 합은 1




- ***Cross-entropy : 확률, 값 -> cost***




- ***One hot***



| Color |
|---|
| Red |
| Red |
| Yellow |
| Green |
| Yellow |

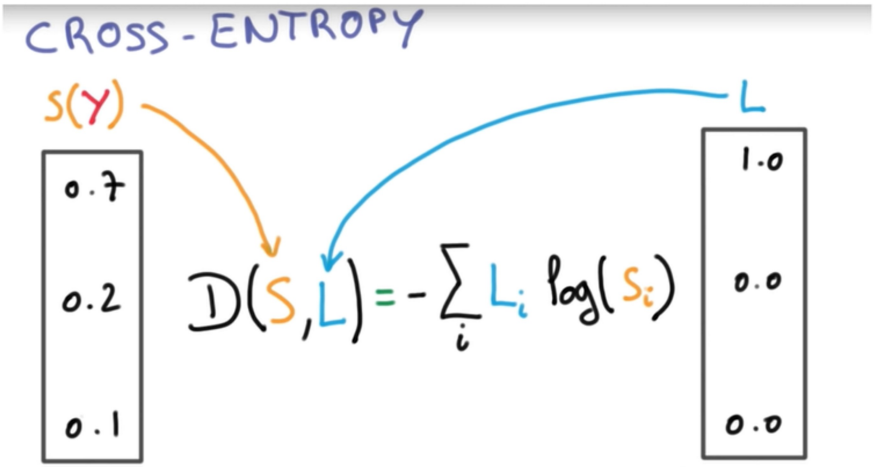
| Red | Yellow | Green |
|---|---|---|
| 1 | 0 | 0 |
| 1 | 0 | 0 |
| 0 | 1 | 0 |
| 0 | 0 | 1 |

2. ***Hypothesis***

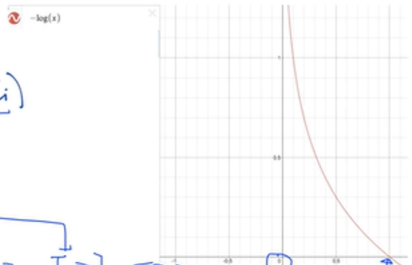
```
logits = tf.matmul(X, W) + b
hypothesis = tf.nn.softmax(logits)

hypothesis = tf.nn.softmax(tf.matmul(X,W))
```

3. ***Cost function***

# Cost function




# Cross-entropy cost function



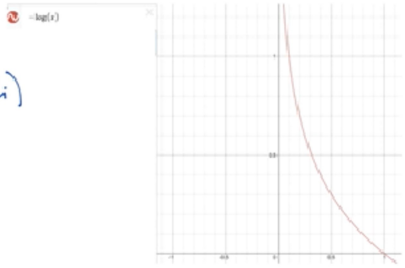
$-\sum_i L_i \log(S_i)$ $\quad -\sum_i L_i \log(\bar{y}_i) = \sum_i L_i * -\log(\bar{y}_i)$

$Y = L = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = B$

$\bar{Y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}_B = B(0K) \Rightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix} \odot -\log\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \odot \begin{bmatrix} \infty \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow 0$

$\bar{Y} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_A = A(X) \Rightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix} \odot -\log\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \odot \begin{bmatrix} 0 \\ \infty \end{bmatrix} = \begin{bmatrix} 0 \\ \infty \end{bmatrix} \Rightarrow \infty$

# Cross-entropy cost function



$-\sum_i L_i \log(S_i)$ $\quad -\sum_i L_i \log(\bar{y}_i) = \sum_i L_i * -\log(\bar{y}_i)$

$L = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_A = A$

$\bar{Y} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}_A = A(0) \Rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} \odot \begin{bmatrix} 0 \\ \infty \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow 0$

$\bar{Y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}_B = B(X) \Rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix} \odot \begin{bmatrix} \infty \\ 0 \end{bmatrix} = \begin{bmatrix} \infty \\ 0 \end{bmatrix} \Rightarrow \infty$

## Logistic cost VS cross entropy

$$C(H(x), y) = ylog(H(x)) - (1-y)log(1-H(x))$$

$$D(S, L) = -\sum_i L_i \log(S_i)$$

### 4. ***Gradient descent algorithm***