## ML Lec 01 - 기본적인 Machine Learning의 용어와 개념 설명

### ML

- due to limitation of explicit programming(spam filter,automatic driving has too many rules)

- field of study that gives computers the ability to learn without being explicitly programmed (Arthur Samuel(1959)

### 1) Supervised learning

learning with labeled examples - training set



Most common problem type in ML

* Image labeling, Email spam filter, Predicting exam score
* from previous score(1~100) based on time spent -> **regression**
* pass/n-pass -> **binary classification**
* Letter grade(A,B,C,E and F) based on time spent : **multi-label classification**

with Training data set

### 2) Unsupervised learning: un-labeled data

ex) google news grouping / word clustering

## 

## 

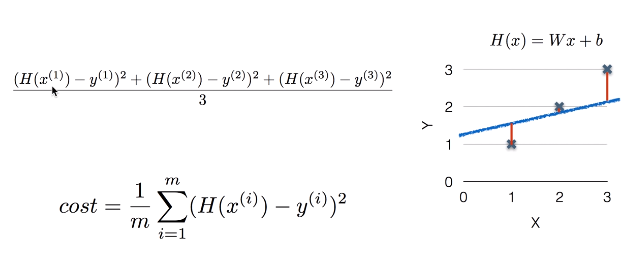
## 02) Linear Regression의 Hypothesis와 cost

### regression: ex) predicting exam score

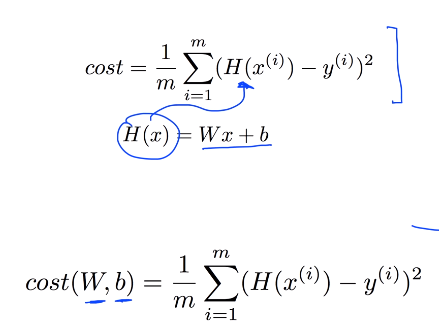
### Hypothesis: H(x) = Wx+b

### Cost function = Loss function

모델이 얼마나 실제에 가까운가? How fit the line to our data?



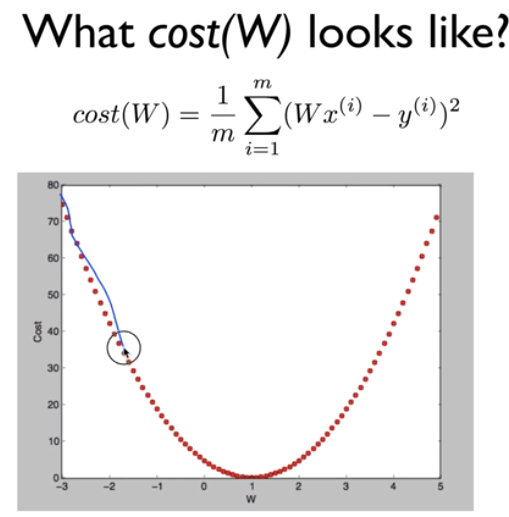
cost function을 아래와 같이 정리할 수 있는데, 학습을 한다는 것은 minimize cost(W,b)한 W와 b를 찾는 것이다.



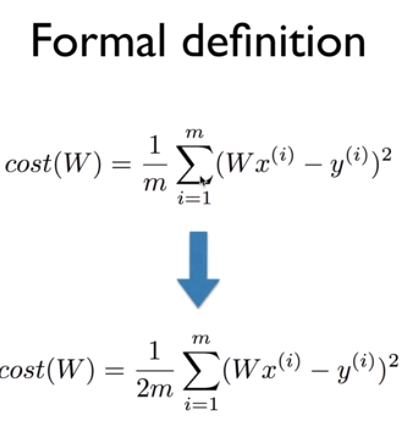
## 03) Linear Regressioin의 최소화 알고리즘의 원리 설명

### Gradient descent algorithm

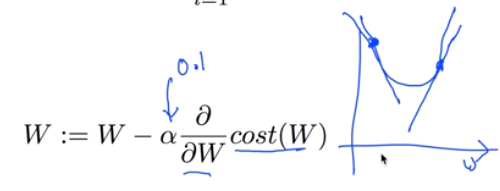
cost function이 최소화되는 점은 기울기가 최소화된 지점이다.



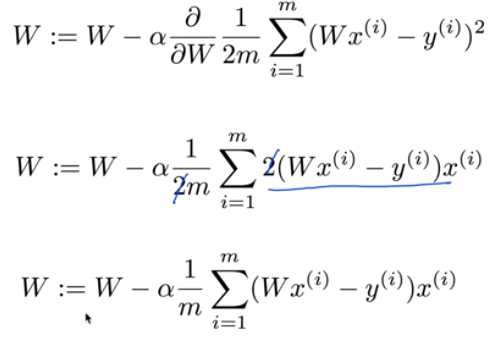
cost function을 미분한다.



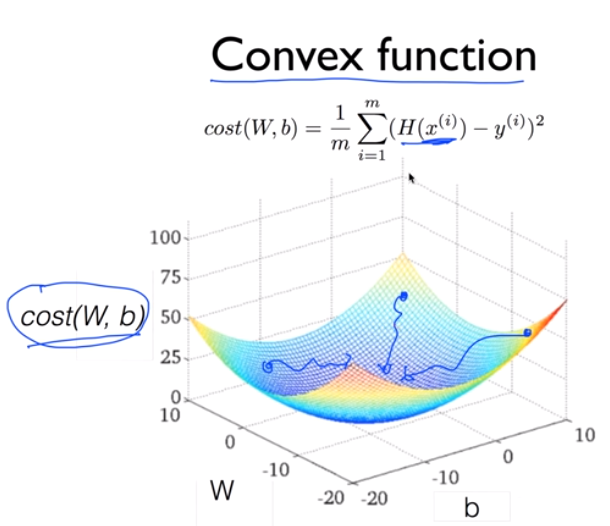
cost(w)를 미분했을 때 양/음 값이 나오면 w값을 -/+조절해본다.



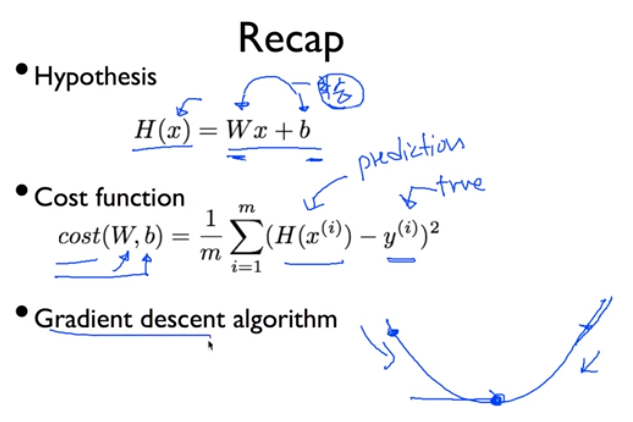
미분식 풀이



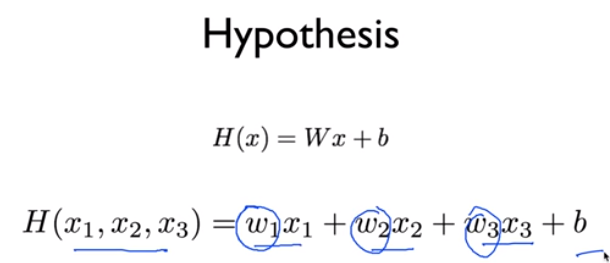
이때 주의: cost(W,b)가 convex function모양을 하고 있어야 성립됨



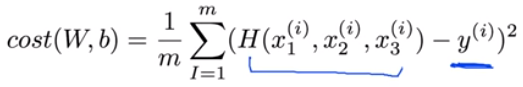
## 04) multi-variable linear regression



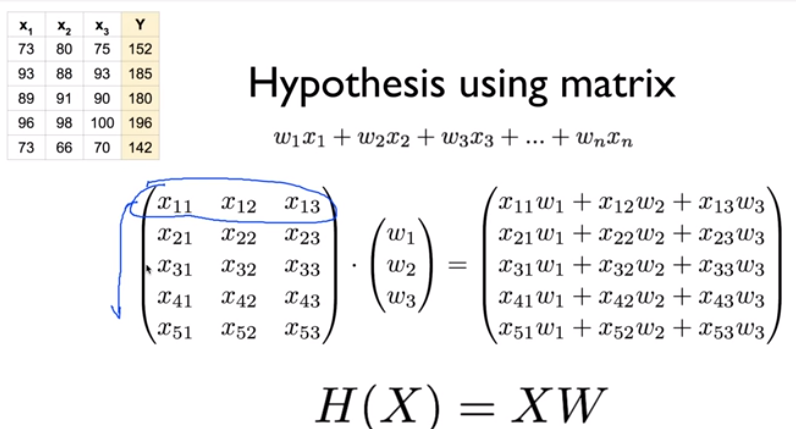
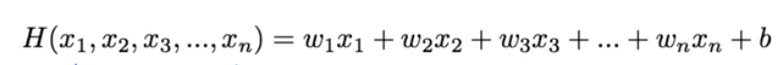
### regression using three inputs(x1,x2,x3)의 Hypothesis



### Cost function



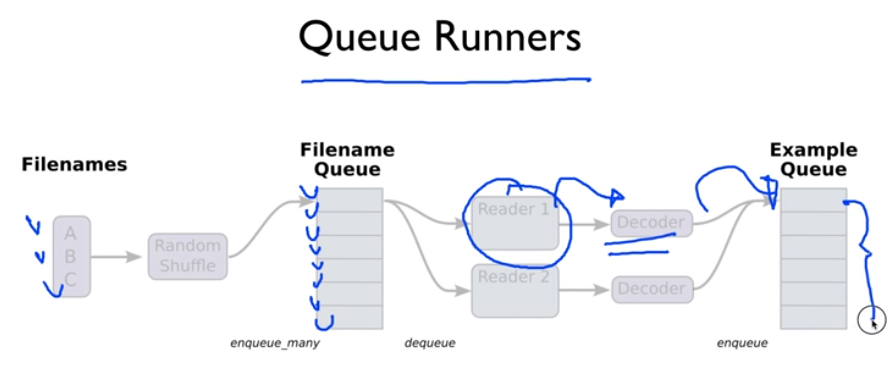
### x의 갯수가 증가한다면? -> Matrix



X의 [a,b] -> a: instance 수 / b: variable 수

H(x)의 [a,b] -> a: instance 수 / b: y값

### Queue Runners



## 05-1) Logistic Classification의 가설 함수 정의

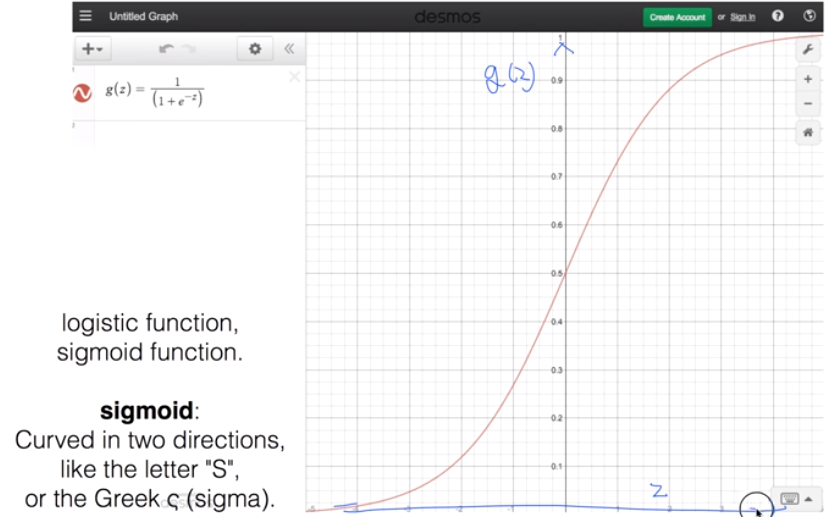
### Binary Classification

0,1 encoding으로 나타낸다

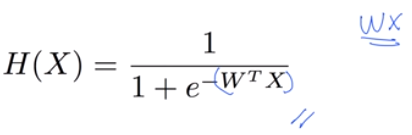
spam(1) or ham(0) / facebook: show(1) or hide(0) / …

기존 Linear regression의 hypothesis에서는 값이 0이하 or 1이상의 수가 나온다.

but binary classification에서는 0 ~ 1만 나와야 한다. => **sigmoid = logistic function**



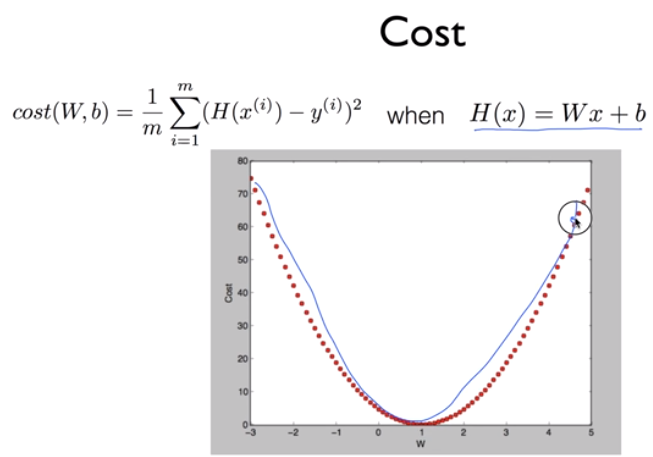
### Logistic Hypothesis



## 

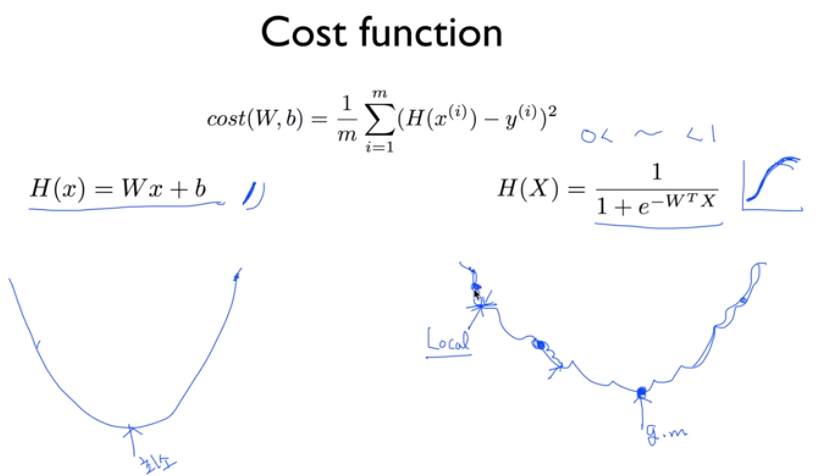
## 05-2) Logistic Regression의 cost 함수 설명

### 기존 linear regression의 cost function



### 기존 Logistic regression cost function의 문제점

기존 linear regression의 cost function은 이차방정식처럼 기울기=0인 지점이 하나이지만,



새로운 logistic hypothesis의 cost function을 그리게 되면 울퉁불퉁해진다.

우리의 목표인 global minimum이 아닌, local minimum에서 멈추게 되므로 새로운 cost function 식이 필요하다.

### 새로운 cost function 이론

classification에서 y의 값은 0과 1밖에 없으므로 y=0일때, y=1일때로 나누어서 생각한다.

1) y=0일때,

H(x): 0이라면, cost값은 0가 나와야 한다.

H(x): 1이라면, cost값은 무한대로 커질 수밖에 없다.

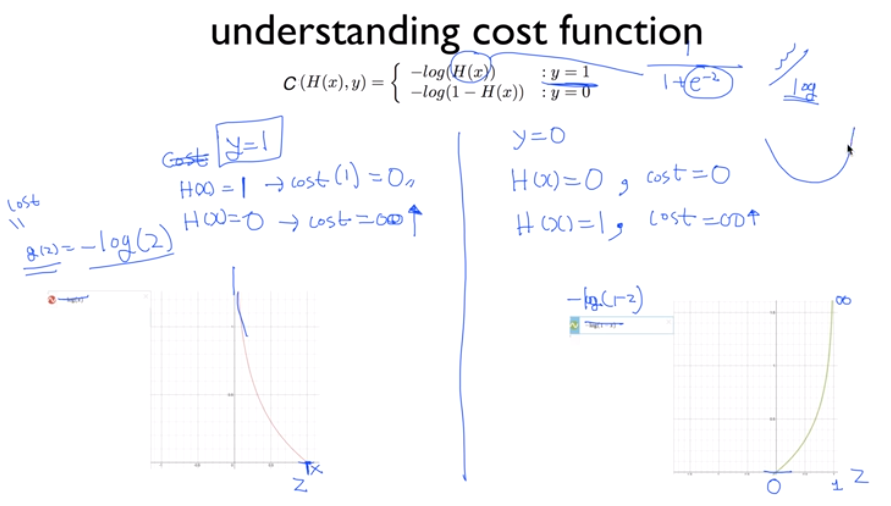
2) y=1일때,

H(x): 1이라면, cost = 0

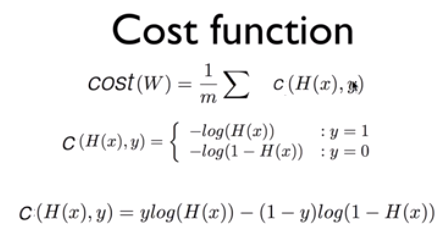
H(x): 0이라면, cost는 무한대로 커진다.

이를 정리하면 아래와 같은 그림과 식이 나온다.

아래 두 그래프를 합치면 linear regression의 cost function함수 모양(아래볼록)이 나옴



### cost function



### Minimize cost - Gradient descent algorithm

