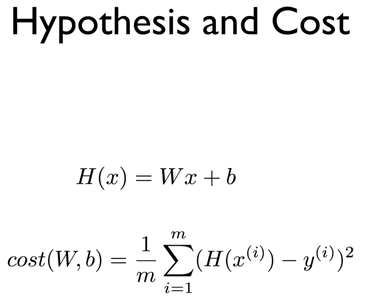
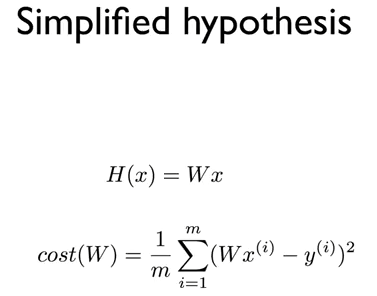
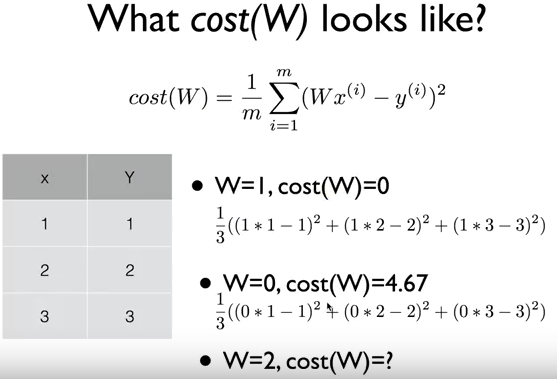
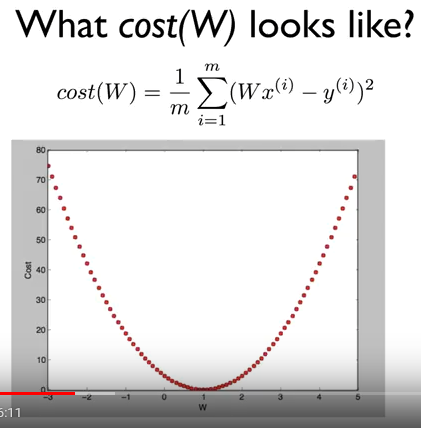
**3장. Linear Regression의 cost 최소화 알고리즘의 원리 설명**



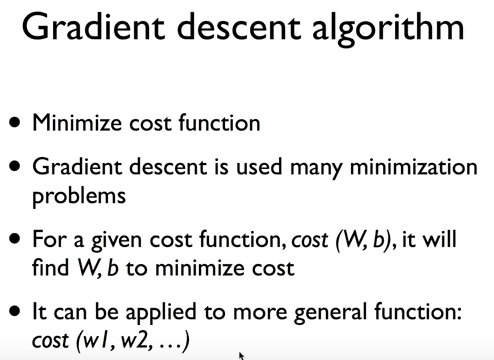
쉽게 설명하기 위하여 b=0 이라고 가정하고, W(weight)를 구하는 문제로 단순화시킴

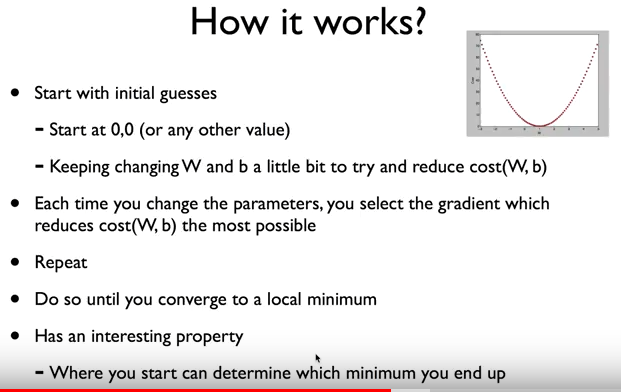




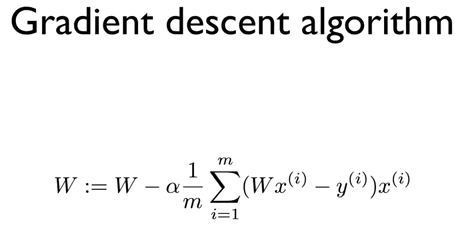


Cost와 W의 그래프에서 cost가 최소가 되는 W를 찾는 방법으로 Gradient descent 알고리즘을 사용한다.

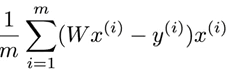




(\*) 아무점(W)에서 시작하며, 해당 점의 기울기를 계산하고 W에서 기울기를 차감하여 새로운 W를 계산한다. 위의 계산을 반복하여 cost가 최저가 되는 W값을 찾는다



(\*) 한 점의 기울기를 구하는 공식은

 이고,

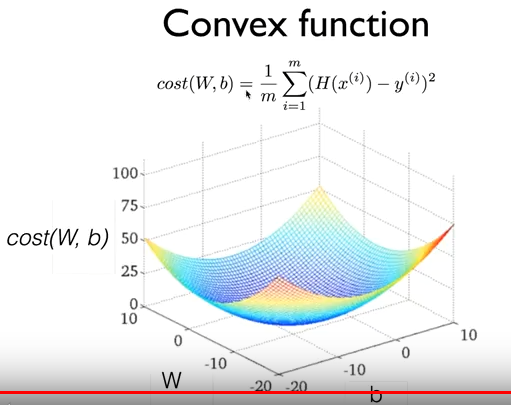
위의 수식에서 알파는 “Learning rate”라고 부르며, 이동시킬 W값의 step size를 나타낸다.

Learning rate값이 너무 크면, cost값이 최저가 되는 W를 찾지 못하게 되고,

너무 작으면, cost값이 최저가 되는 W를 찾기까지 너무 많은 시간이 소요된다.

(일반적으로는 0.1 정도로 학습한다)

Gradient descent 알고리즘을 사용하려면, cost(W,b)와 W, b 로 3차원그래프를 그렸을 때 아래와 같은 모양이 되어야 한다. 즉, Cost(W,b)함수가 Convex function 인 경우만 사용 가능하다.



아래와 같이 cost함수가 convex function이 아닌 경우는 시작점에 따라서 최저점이 다르게 계산되므로 Gradient descent 알고리즘을 사용할 수 없다.

