

Gradient Descent (경사 하강법)

<https://www.youtube.com/watch?v=IHZwWFHWa-w&t=423s> :10분 27

Gradient Descent (경사 하강법)

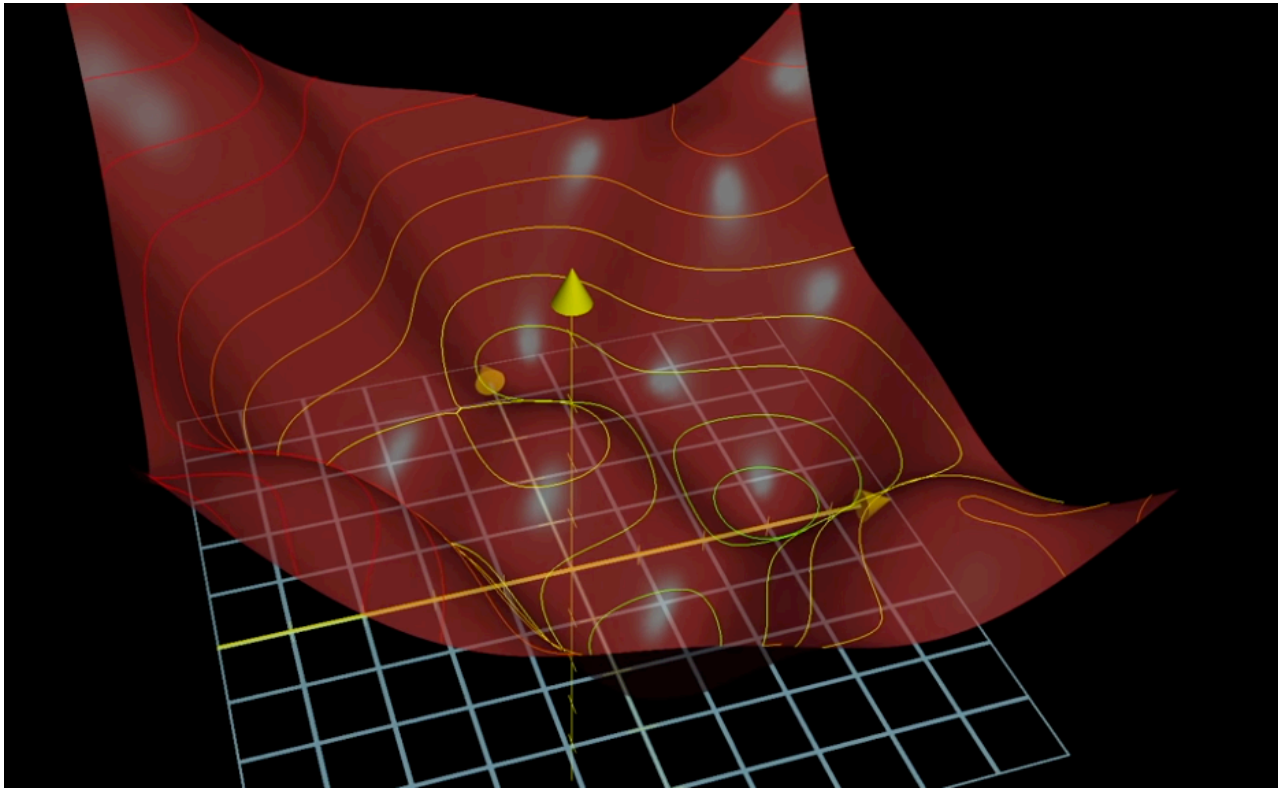
결론

$$Cost = \sum(\text{예측값} - \text{정답})^2$$

Ex) Input : 강아지 사진

	예측값.	정답	
강아지.	0.53.	1	$\Rightarrow (0.53 - 1)^2 + (0.23 - 0)^2 + (0.24 - 0)^2$
고양이.	0.23.	0	
고래.	0.24.	0	

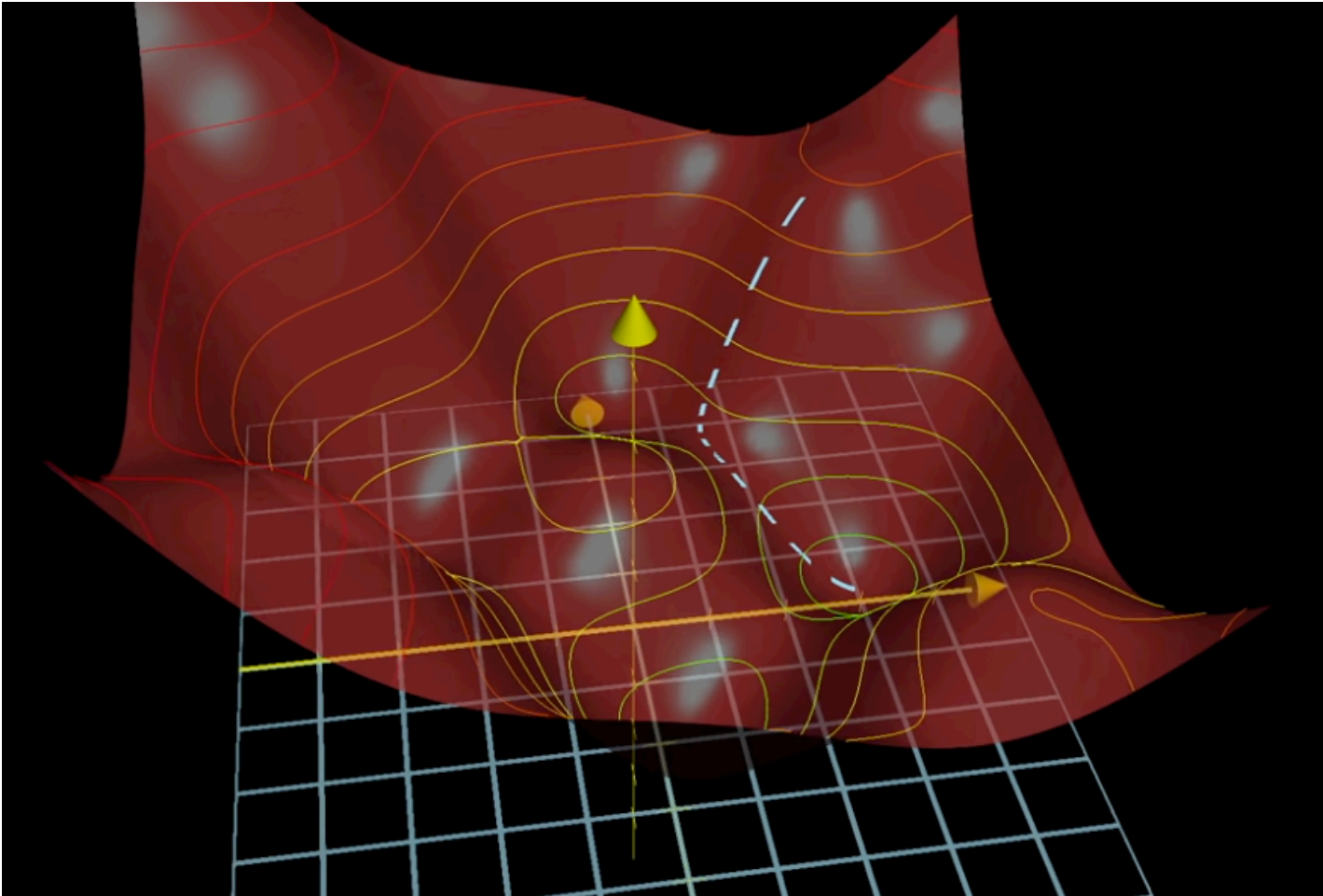
Gradient Descent (경사 하강법)



$C(w_1, w_2)$

x, y 축이 각각 w_1, w_2 이고,
z축이 Cost일 때 이런 모양의
Cost함수가 만들어 진다.

Gradient Descent (경사 하강법)



Minima를 찾기 위해서 가장 쉬운 방법은 Gradient Descent!

W_1, W_2 각 특정 값일 때 Cost함수의 미분 값을 계산하고 그 미분 값의 반대 방향으로 움직여 준다.

Local minima: Doable
Global minima: Very hard

Gradient Descent (경사 하강법)

이렇게 Global Minima를 찾는 방법에 여러가지 방법이 있는데, 최적 점을 찾는다고 해서 이 방법을 Optimizer라고 한다.

Gradient Descent 방법은 가장 기초적으로 사용되는 Optimizer.

최근에는 Adam Optimizer라는 Gradient Descent를 의용한 최적화 방법을 많이 사용한다.

<https://ganghee-lee.tistory.com/24>

Gradient Descent (경사 하강법)

추가! Learning rate

Learning rate도 중요한 parameter 중 하나.

Learning rate가 너무 크면, minima를 지나칠 수도 있고, 너무 작으면, local minima에 들어갈 확률이 커지고, 학습이 느리다.

요즘에는 framework가 Learning rate를 자동으로 조절해준다.

