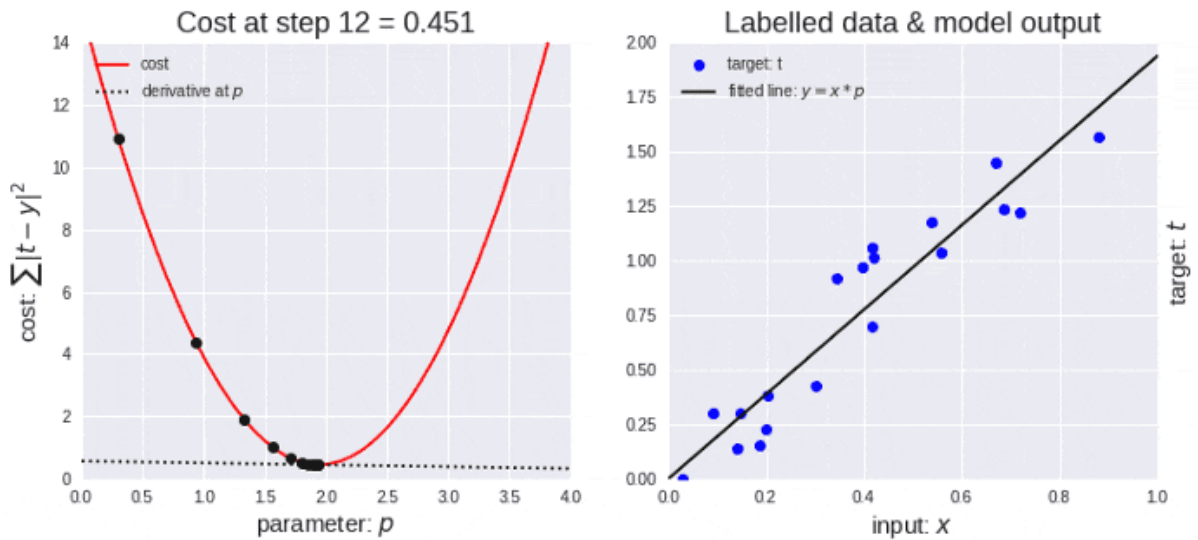


▶ 알고리즘이 **iterative**하다는 것: **gradient descent**와 같이 결과를 내기 위해서 여러 번의 최적화 과정을 거쳐야 되는 알고리즘



optimization 과정

▶ 다루어야 할 데이터가 너무 많기도 하고(메모리가 부족하기도 하고) 한 번의 계산으로 최적화된 값을 찾는 것은 힘듭니다. 따라서, 머신 러닝에서 최적화(optimization)를 할 때는 일반적으로 여러 번 학습 과정을 거칩니다. 또한, 한 번의 학습 과정 역시 사용하는 데이터를 나누는 방식으로 세분화 시킵니다.

이때, **epoch**, **batch size**, **iteration**라는 개념이 필요합니다.

### - epoch

*One Epoch is when an ENTIRE dataset is passed forward and backward through the neural network only ONCE*

\*\*\*

(한 번의 **epoch**는 인공 신경망에서 전체 데이터 셋에 대해 **forward pass/backward pass** 과정을 거친 것을 말함. 즉, 전체 데이터 셋에 대해 한 번 학습을 완료한 상태)

\*\*\*

▶ 신경망에서 사용되는 역전파 알고리즘(**backpropagation algorithm**)은 파라미터를 사용하여 입력부터 출력까지의 각 계층의 **weight**를 계산하는 과정을 거치는 순방향 패스(**forward pass**), **forward pass**를 반대로 거슬러 올라가며 다시 한 번 계산 과정을 거쳐 기존의 **weight**를 수정하는 역방향 패스(**backward pass**)로 나뉩니다. 이 전체 데이터 셋에 대해 해당 과정 (**forward pass + backward pass**)이 완료되면 한 번의 **epoch**가 진행됐다고 볼 수 있습니다.

역전파 알고리즘이 무엇인지 잘 모른다고 하더라도 **epoch**를 전체 데이터 셋에 대해 한 번의 학습 과정이 완료됐다고 단편적으로 이해하셔도 모델을 학습 시키는 데는 무리가 없습니다.

epochs = 40이라면 전체 데이터를 40번 사용해서 학습을 거치는 것입니다.

---

▶ 우리는 모델을 만들 때 적절한 epoch 값을 설정해야만 underfitting과 overfitting을 방지할 수 있습니다.

epoch 값이 너무 작다면 underfitting이 너무 크다면 overfitting이 발생할 확률이 높은 것이죠.

- batch size

Total number of training examples present in a single batch.

\*\*

- iteration

The number of passes to complete one epoch.

\*\*

batch size는 한 번의 batch마다 주는 데이터 샘플의 size. 여기서 batch(보통 mini-batch라고 표현)는 나뉜 데이터 셋을 뜻하며 iteration는 epoch를 나누어서 실행하는 횟수라고 생각하면 됨.

\*\*

▶ 메모리의 한계와 속도 저하 때문에 대부분의 경우에는 한 번의 epoch에서 모든 데이터를 한꺼번에 집어넣을 수는 없습니다. 그래서 데이터를 나누어서 주게 되는데 이때 몇 번 나누어서 주는가를 iteration, 각 iteration마다 주는 데이터 사이즈를 batch size라고 합니다.



1 Epoch : 모든 데이터 셋을 한 번 학습

1 iteration : 1회 학습

minibatch : 데이터 셋을 batch size 크기로 쪼개서 학습

ex) 총 데이터가 100개, batch size가 10이면,

1 iteration = 10개 데이터에 대해서 학습

1 Epoch = 100/batch size = 10 iteration

- 정리

---

전체 2000 개의 데이터가 있고, epochs = 20, batch\_size = 500이라고 가정합니다.

---

그렇다면 **1 epoch**는 각 데이터의 **size**가 **500**인 **batch**가 들어간 네 번의 **iteration**으로 나누어집니다.

그리고 전체 데이터셋에 대해서는 **20** 번의 학습이 이루어졌으며, **iteration** 기준으로 보자면 총 **80** 번의 학습이 이루어진 것입니다.

---