

NN adv 6주차 (learning rate scheduler 정리) 1-1

- pytorch에서 제공하는 learning rate scheduler 소개하고자 한다.

1. LambaLR

- 가장 유연한 scheduler 중 하나로, 어떻게 scheduling을 이뤄낼지에 대해서 lambda함수를 이용하거나, 간단한 지수함수와 같은 것을 이용해서 정한다.
- 사용되는 파라미터는 `optimizer`, `lr_lambda` 이다.
- 코드 예시

```
scheduler = LambdaLR(optimizer, lr_lambda = lambda epoch: 0.95 ** epoch)
```

→ `lambda` 함수를 이용하여 0.95^{epoch} 형태로 나타냄

2. Custom CosineAnnealingWarmRestarts

- CosineAnnealingLR의 변형된 형태로, 현재 많은 논문에서도 사용 중이고, SGDR로 알려져 있다.
- 특히 논문 bag of tricks for image classification에서 사용한 최적화 방법으로 좋은 성능을 보였다.
- 사용되는 파라미터는 `T_0`, `T_mult`, `eta_max` 외에 `T_up`, `gamma` 이다.
- `T_0`: 첫 번째 주기의 길이, `T_mult`: 주기의 길이를 조절, `eta_max`: 학습률이 증가하는 최대값 `T_up`: 각 주기 이후에 학습률을 증가시키는 시간 `gamma`: 주기의 길이를 늘릴 때 곱해지는 값
- 코드 예시

```
optimizer = optim.Adam(model.parameters(), lr = 0)
```

```
scheduler = CosineAnnealingWarmRestarts(optimizer, T_0 = 150, T_Mult = 1, eta_max = 0.1, T_up = 10, gamma = 0.5)
```

→ `T_0=150 epoch`의 초기 주기값 후에 다시 0으로 줄어들게 된다. 이 때, `T_up=10 epoch` 만에 learning rate는 0 → `eta_max` 까지 증가하게 되고 다음 주기에 는 `gamma=0.5` 만큼 곱해진 `eta_max * gamma` 만큼 warm up start 하여 learning rate가 증가하게 된다.

▼ 참고자료

