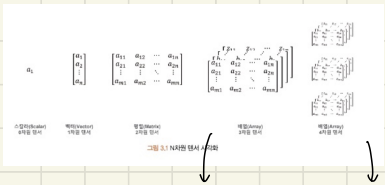
03 파이토치 기초

텐서

≒ 넘파이. 차이점은 cpu-gpu 사용함



N, M, K

이미지의 경우 C, H, W

N, C, H, W (이미지의 개수)

텐서 생성

torch.tensor( ) : 입력된 데이터를 복사해 텐서로 변환. 데이터 잃어버림

torch.Tensor( ) : 텐서인스턴스를 생성하는 함수. 데이터 잃어도 됨.

텐서 속성

형태(shape), 자료형(dtype), 장치(device) ← gpu 가속연산

차원 변환

예제 3.3 텐서 차원 변환

```
import torch

tensor = torch.rand(1, 2)
print(tensor)
print(tensor.shape)

tensor = tensor.reshape(2, 1)
print(tensor)
print(tensor.shape)
```

출력 결과

```
tensor([[0.6499, 0.3419]])
torch.Size([1, 2])
tensor([[0.6499],
        [0.3419]])
torch.Size([2, 1])
```

장치 설정 ~ 변환 → 헷갈리지 모르겠당!

손실함수

손실함수 ⊂ 비용함수 ⊂ 목적함수

: 단일샘플의 실젯값과 예측값의 차이가 발생했을때 오차가 얼마인지 계산하는 함수

EX) 생활회피에서의 손실함수 (연속형 변수)

$$SE = (Y_i - \hat{Y}_i)^2$$

$$SSE = \sum_{i=1} (Y_i - \hat{Y}_i)^2$$

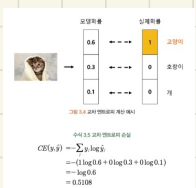
$$MSE = \frac{1}{n}\sum_{i=1} (Y_i - \hat{Y}_i)^2$$

$$RMSE = \sqrt{MSE}$$ → 루트를 씌워서 제곱으로 인해 발생한 오차를 감소

교차 엔트로피 (이산형 변수)

: 실젯값의 확률분포와 예측값의 확률분포 차이를 계산

$$CE(y, \hat{y}) = -\sum_i y_i \log \hat{y}_i$$



수식 3.5 교차 엔트로피 손실

$$CE(y,\hat{y}) = -\sum_i y_i \log \hat{y}_i$$
$$= -(1\log 0.6 + 0\log 0.3 + 0\log 0.1)$$
$$= -\log 0.6$$
$$= 0.5108$$

최적화

: 목적함수의 결괏값을 최적화하는 방법을 찾는 알고리즘

경사하강법

↳ 함수의 기울기가 낮은 곳으로 계속 이동시켜 극값에 도달할 때까지 반복하는 알고리즘이다.

W_0 = 임의의 value (임의의 초깃값)

$W_{t+1} = W_t - \alpha \nabla f(W_t)$ 기울기?

↳ step size를 결정

수식 3.8 가중치 경신 방법

$$\begin{aligned} W_{t+1} &= W_t - \alpha\frac{\partial}{\partial W} MSE(W,b) \\ &= W_t - \alpha\frac{\partial}{\partial W}\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \\ &= W_t - \alpha\frac{\partial}{\partial W}\sum_{i=1}^{n}\left[\frac{1}{n}(Y_i - (W_t \times x + b_t))^2\right] \\ &= W_t - \alpha \times \frac{2}{n}\sum_{i=1}^{n}[Y_i - (W_t \times x + b_t) \times (-x)] \\ &= W_t - \alpha \times \frac{2}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i) \times (-x) \\ &= W_t - \alpha \times \frac{2}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i) \times x \\ &= W_t - \alpha \times 2E[(\hat{Y}_i - Y_i) \times x] \end{aligned}$$

수식 3.9 가중치 경신 방법 일반형

$$W_{t+1} = W_t - \alpha \times E[(\hat{Y}_i - Y_i) \times x]$$

학습률 (learning rate : α)

↳ 임의의 값

학습률이 너무 높거나 너무 낮으면 여러가지 문제가 발생할 수 있음!

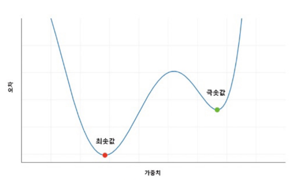


그림 3.11 최솟값과 극솟값

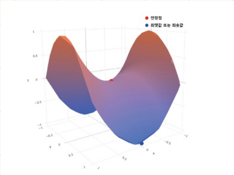


그림 3.12 안장점이 존재하는 그래프

< 여러가지 문제들 >

~79p

넘파이로 하는건

읽고만 넘어감