

5. Maximum A Posterior Estimation

Bayes Theorem

이전 슬라이드 : Likelihood(MLE)/NLL 을 사용해서 적합한 파라미터(seta)를 찾았다면

이번 슬라이드 : Likelihood와 Prior을 이용해서 더 많은 정보로 더 정확하게 파라미터 (seta)를 찾는 것

$$P(h|D) = \frac{\overbrace{P(D|h)}^{\text{Likelihood}} \overbrace{P(h)}^{\text{Prior}}}{\underbrace{P(D)}_{\text{Evidence}}}$$

h : hypothesis
 D : Data

Bayesian vs Frequentist

Bayesian 관점

- 파라미터 또한 random variable이며, prior 분포를 따를 것.
- 미래의 uncertainty까지 고려
- Prior에 대한 가정이 필요

$$\begin{aligned}\hat{\theta} &= \operatorname{argmax}_{\theta \in \Theta} P(\theta|D) \\ &= \operatorname{argmax}_{\theta \in \Theta} \frac{P(D|\theta)P(\theta)}{P(D)} \\ &= \operatorname{argmax}_{\theta \in \Theta} P(D|\theta)P(\theta)\end{aligned}$$

↓
아직 연구 미전함

↓
θ 관한 한층원리를 가지고 있다는 말임

Frequentist 관점

- 파라미터는 최적화의 대상
- 현재까지의 정보를 바탕으로 추정
- Overfitting에 취약함

$$\hat{\theta} = \operatorname{argmax}_{\theta \in \Theta} P(D; \theta)$$

← "Prior" 더 많은 정보를 Bayesian에서는 가지고 있음



