

# 泛化理论

## 第二章 一致收敛

### §2.1.1 Uniform Convergence

@ 滕佳焯

## Recall:

*Generalization*: measuring how model performs on *unseen* data.

*Consistency*: when we have infinite training samples ( $n \rightarrow \infty$ ), the estimator  $\hat{\beta}$  should converge to the true parameter  $\beta^*$ .

*Traditional statistical models*: linear regression, generalized linear models, kernel tricks, empirical process...

## Today's topic:

Uniform convergence (UC): definition and intuition

## Generalization Gap

With good optimization (ERM, recall 0.2.1)

$$\mathcal{L}(\hat{f}) - \inf_{f \in \mathcal{F}} \mathcal{L}(f) = [\mathcal{L}(\hat{f}) - \hat{\mathcal{L}}(\hat{f})] + [\hat{\mathcal{L}}(\hat{f}) - \hat{\mathcal{L}}(f^*)] + [\hat{\mathcal{L}}(f^*) - \mathcal{L}(f^*)]$$

Generalization Gap

ERM,  $\leq 0$

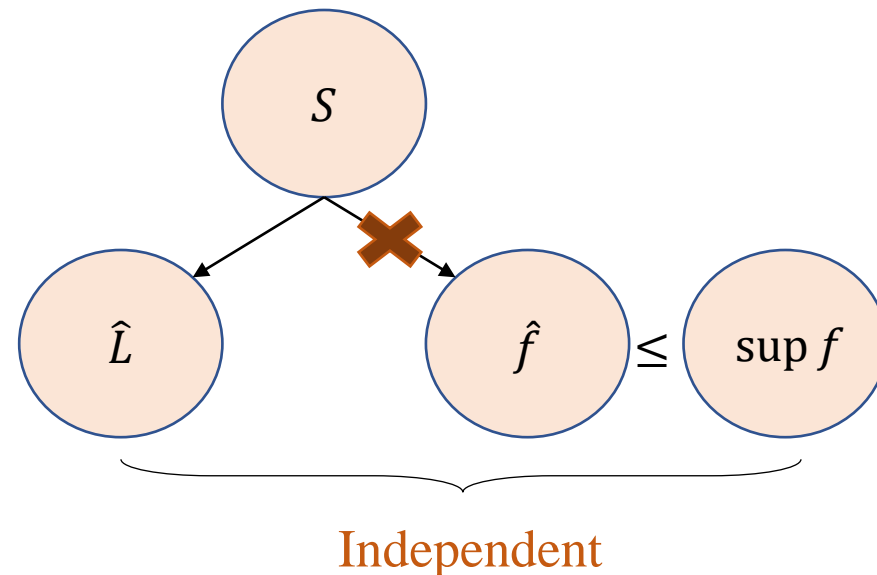
Concentration

Key difficulty: **dependency** between training loss  $\hat{L}$  and trained parameter  $\hat{f}$  (via training set  $S$ )

- **Decouple** the dependency via *uniform convergence*:

$$L(\hat{f}) - \hat{L}(\hat{f}) \leq \sup_{f \in \mathcal{F}} |L(f) - \hat{L}(f)|.$$

Decouple:  $\mathcal{F}$  should not depend on training set  $S$ .  
 $\mathcal{F}$ : hypothesis class.



## More intuition on Uniform Convergence

Key difficulty: **dependency** between training loss  $\hat{L}$  and trained parameter  $\hat{f}_{erm}$  (via training set  $S$ )

$$L(\hat{f}) - \hat{L}(\hat{f}) \leq \sup_{f \in \mathcal{F}} |L(f) - \hat{L}(f)|.$$

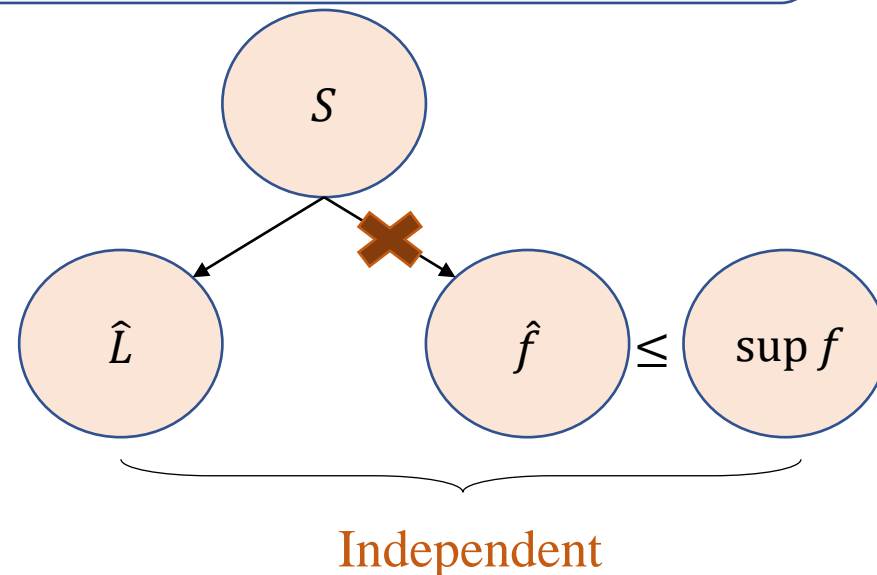
Why we hate dependency?

- It is hard to analyze...
- It violate the requirement of concentration...

Is this relaxation too loose?

- Yes... But we have no other choice ☹️

Techniques to deal with UC: VC dimension, Rademacher complexity ...



## Take-away messages

- (a) Difficulty in generalization gap: dependency
- (b) How to deal with the dependency? Uniform convergence!
- (c) How to do uniform convergence? Taking sup over the parameter space (hypothesis class).

All the slides will be available at [www.tengjiaye.com/generalization](http://www.tengjiaye.com/generalization) soon.

@ 滕佳烨

Thanks!