2019.1.7

Natural of this course

less structured

like a seminar

Recent advance

方式：

阅读论文 → 讨论

What's the course about.

Integration of opt. and stat. techniqus

Case study: Statistical estimation problems

- Samples: $z_1, z_2, \cdots z_n$

n个样本，

- parameters: $\theta_1, \theta_2, \cdots, \theta_d$

d个参数.

e.g.: linear regression

$y = X\theta^* + \varepsilon$ $\quad z = (X, y)$ samples

$y \in \mathbb{R}^n$, $\theta^* \in \mathbb{R}^d$, $\varepsilon \in \mathbb{R}^n$ noise

θ^* : star means ground truth

Classical Setting:

$n >> d$ (Overdetermined) 通常假设 n 远大于 d

$\hat{\theta} = \arg\min_{\theta \in \mathbb{R}^d} \|y - X\theta\|_2^2$ Least-square estimator

θ^* 是 ground truth (We don't know)

$\hat{\theta}$ 是 optimal solution to the opt. problem

called estimator.

Note: $\hat{\theta} \neq \theta^*$ in general

find the error bound between

$\hat{\theta}$ and θ^*

From Stat. we know:

If $\varepsilon \sim N(0, \sigma^2)$, Then the

maximum likelihood estimator (MLE)
is given by $\hat{\theta}$ (LS解析解)

Issues/Observations:

① Optimization side: