3. Recall the Ideas for M-estimation
Recall: Ideas for M-estimation

Start of transcript. Skip to the end.

(Caption will be displayed when you start playing the video.)
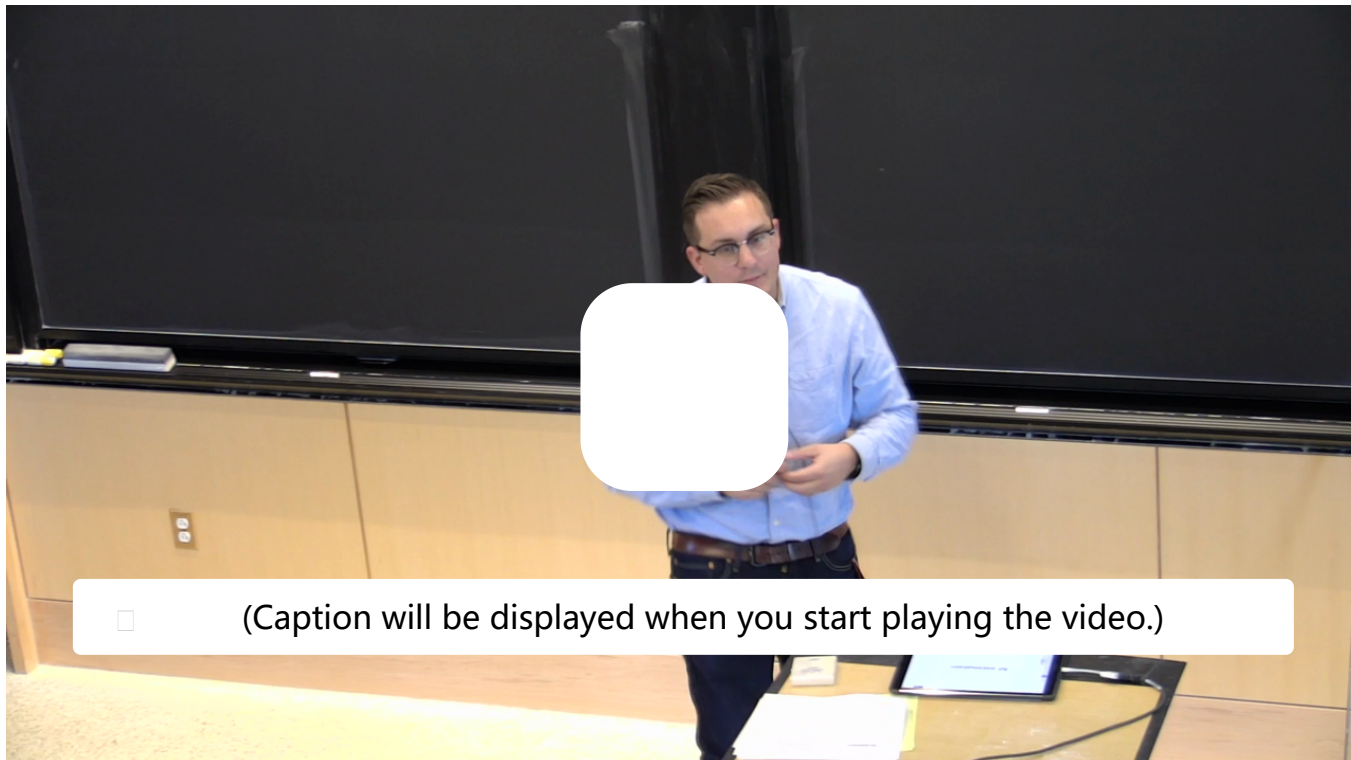
So we're concluding this chapter talking about M-estimation.

And M-estimation is essentially a generalization

of the main strategy that we had when

we built the maximum likelihood estimator.

So what was the MLE strategy?

Well, it said if you look at the KL divergence between P theta

star and P theta as a function of theta, then

视频
下载视频文件

字幕
下载 SubRip (.srt) file
下载 Text (.txt) file

Concept check: Defining M-estimators

1/1 point (graded)

Suppose we have access to a distribution $\mathbf{P}$ which has an unknown parameter μ^* that we would like to estimate from samples $X_1, \ldots, X_n \overset{iid}{\sim} \mathbf{P}$. Suppose we have a **loss function** $\rho(x, \mu)$ with the property that

$$\mu^* = \operatorname{argmin}_{\mu \in \mathbb{R}} \mathbb{E}_{X \sim \mathbf{P}}\left[\rho(X, \mu)\right].$$

What commonly used statistical trick is used to define an M-estimator? (Refer to the slides.)

○ Using the KL divergence instead of TV distance.

○ The method of moments.

◉ Replacing expectations with averages. ☐

Solution:

The correct response is "Replacing expectations with averages." Indeed, we have that the equation

$$\mu^* = \operatorname{argmin}_{\mu \in \mathbb{R}} \mathbb{E}_{X \sim \mathbf{P}}\left[\rho(X, \mu)\right]$$

becomes

$$\widehat{\mu} = \text{argmin}_{\mu \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^{n} [\rho(X_i, \mu)]$$

upon replacing the expectation by an average over the sample. Here, $\widehat{\mu}$ is precisely the M-estimator associated with $\rho(x, \mu)$.

The response "Using the KL divergence instead of TV distance." is incorrect. Rather, the KL divergence was used specifically in the context of maximum likelihood estimation. It does not play a role in the context of M-estimation.

The response "The method of moments." is also incorrect. The method of moments is a tool for parameter estimation which is distinct from M-estimation. The method of moments is not what is used to define an M-estimator.

提交   你已经尝试了1次（总共可以尝试2次）

☐   Answers are displayed within the problem

讨论

显示讨论

主题： Unit 3 Methods of Estimation:Lecture 12: M-Estimation / 3. Recall the Ideas for M-estimation