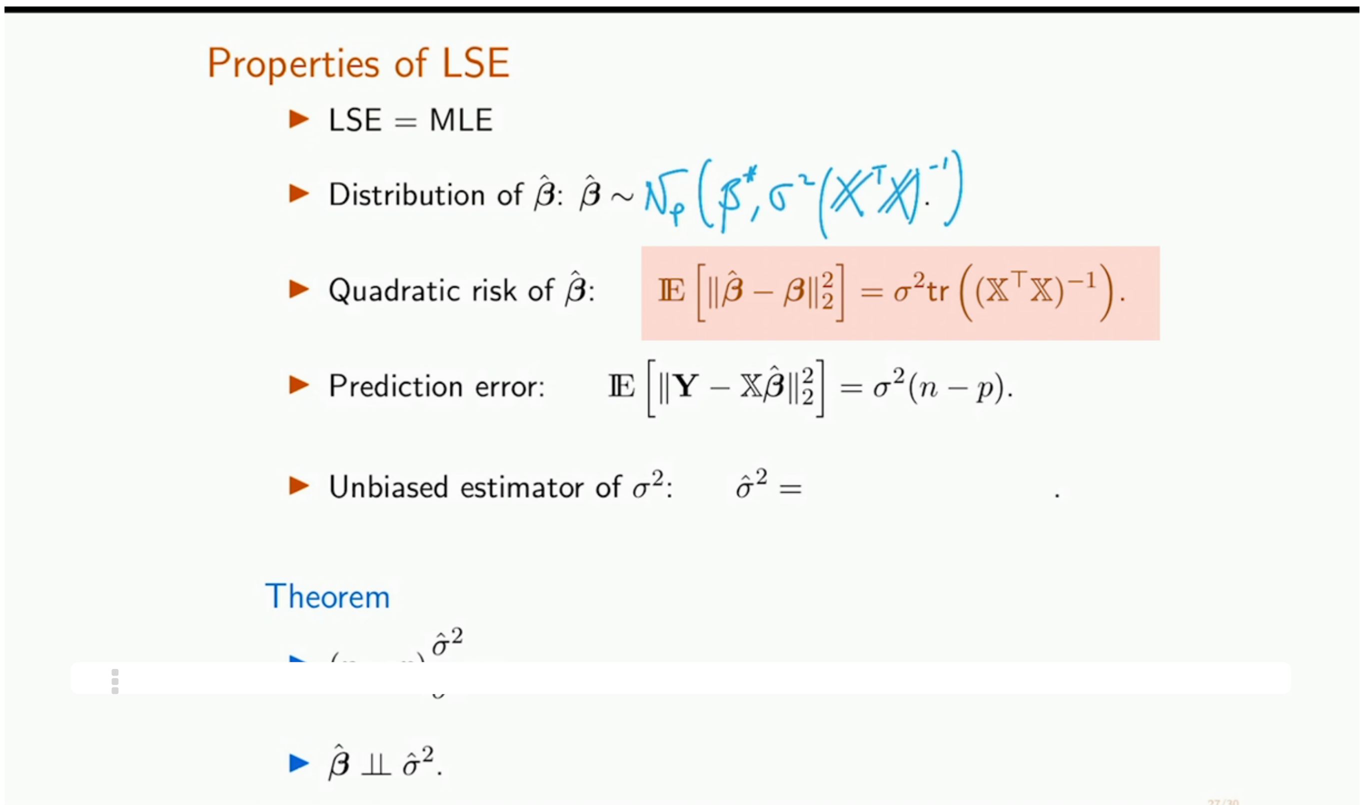
9. Quadratic Risk and Variance
Quadratic Risk



0:00 / 0:00 1.0x

Start of transcript. Skip to the end.

So now let's compute the quadratic risk of beta hat.

So you can see now that there's the trace coming in here, so--

I'm sorry, this is becoming heavier and heavier

in linear algebra.

So I'm going to keep my formula for beta hat.

So what is beta hat minus beta star?

Well, we know that this is a Gaussian

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

The Quadratic Risk

1/1 point (graded)

We would like to analyze the *typical* *error in $\hat{\beta}$ compared to the true parameter, β*, which we may define as $\mathbb{E}\left[\|\hat{\beta} - \beta\|_2^2\right]$. On the other hand, we might also consider the *typical* *error between the predictions $\mathbb{X}\hat{\beta}$ and the observations $\mathbf{Y}$*, which we may define as $\mathbb{E}\left[\|\mathbf{Y} - \mathbb{X}\hat{\beta}\|_2^2\right]$.

These are respectively called the **quadratic risk of $\hat{\beta}$** and the **prediction error** . (The prediction error will be discussed in the next video.)

What happens to these errors as σ^2 increases?

- ( ) The error in $\hat{\beta}$ increases, but the prediction error decreases.
- ( ) The error in $\hat{\beta}$ decreases, but the prediction error increases.
- ( ) Both errors decrease.
- (x) Both errors increase. ✔

Solution:

σ^2 is the variance of each coordinate of ϵ. As the σ^2 increases, the data becomes more noisy. In particular, the task of estimating $\hat{\boldsymbol{\beta}}$ ought to become harder, and it is intuitive that $\mathbf{Y}$ becomes further from the prediction $\mathbb{X}\hat{\boldsymbol{\beta}}$. To make this concrete, recall the following formulas, which hold in the homoscedastic Gaussian case:

$$\mathbb{E}[\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\|_2^2 = \sigma^2 \mathbf{tr}\left((\mathbb{X}^T\mathbb{X})^{-1}\right)$$

$$\mathbb{E}\left[\|\mathbf{Y} - \mathbb{X}\hat{\boldsymbol{\beta}}\|_2^2\right] = \sigma^2\,(n-p)$$

(In our scenario, $n = 1000$ and $p = 2$.)

Submit

You have used 1 of 3 attempts

Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 6 Linear Regression:Lecture 20: Linear Regression 2 / 9. Quadratic Risk and Variance