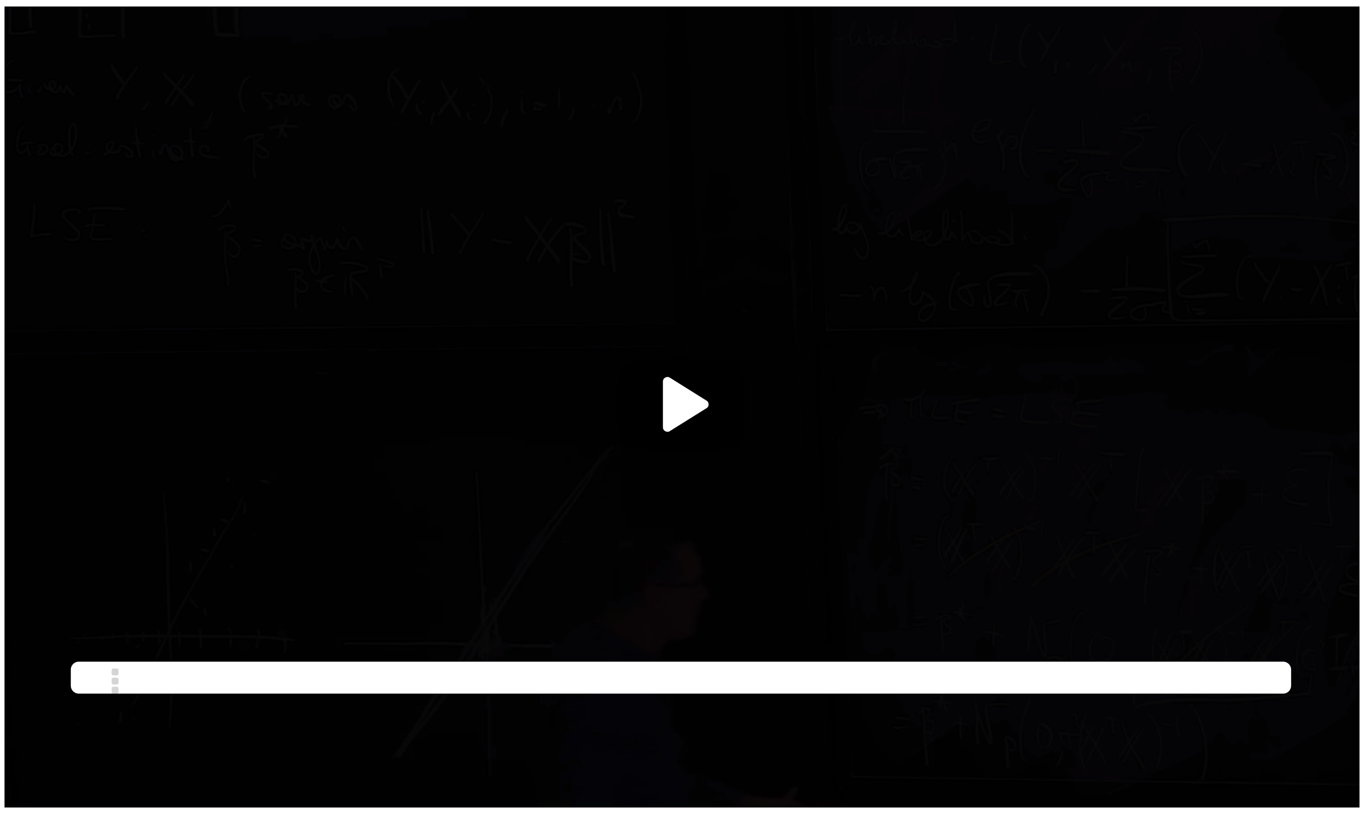
7. Distribution of the Least Square Estimator

Distribution of the Least Square Estimator



8:28 / 8:28 1.0x

unstable.

OK, so this is what this thing is reflecting.

The first product-- the first term in the product sigma

squared is telling you how far those points are.

And the second one is really telling you how much leverage

you get from here.

So this product is really an aspect ratio

of what your scatterplot looks like.

End of transcript. Skip to the start.

Video
Download video file

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

Gaussian Noise

3/3 points (graded)

Recall that the Least-Squares Estimator $\hat{\boldsymbol{\beta}}$ has the formula

$$\hat{\boldsymbol{\beta}} = \left(\mathbb{X}^T\mathbb{X}\right)^{-1}\mathbb{X}^T\mathbf{Y}.$$

If we assume that the vector $\boldsymbol{\varepsilon}$ is an n-dimensional Gaussian with mean 0 and covariance $\sigma^2 I_n$ for some known $\sigma^2 > 0$, then:

“The model is **homoscedastic** ; i.e. $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d."

- (•) True ✔
- ( ) False

“In the deterministic design setting, the LSE $\hat{\boldsymbol{\beta}}$ is a Gaussian random variable."

- (•) True ✔
- ( ) False

“If $\mathbb{X}$ is a random variable, then the LSE $\hat{\boldsymbol{\beta}}$ is still a Gaussian random variable."

True

False ✔

Solution:

- "**The model is homoscedastic ; i.e. $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d.**" is true. The covariance matrix is a diagonal matrix. Recall the useful fact that the ith coordinate of a multi-dimensional gaussian is also a gaussian. In the case where the covariance matrix is diagonal, the coordinates also happen to be independent. Therefore, the first statement is **true**.

- "**In the deterministic design setting, the LSE $\hat{\beta}$ is a Gaussian random variable**" is true. We have

$$\hat{\beta} = (\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\mathbf{Y} = \beta + (\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\epsilon.$$

  By using the result of the first exercise, we arrive at the conclusion that $\hat{\beta}$ is also Gaussian: $\hat{\beta} \sim \mathcal{N}\left(\beta, \sigma^2(\mathbb{X}^T\mathbb{X})^{-1}\right).$

- "**If $\mathbb{X}$ is a random variable, then the LSE $\hat{\beta}$ is still a Gaussian random variable**" is false. *The assumption that $\mathbb{X}$ is deterministic/constant is crucial.* If $\mathbb{X}$ were a generic random variable, then $\hat{\beta}$ might no longer be Gaussian.

  Perhaps the simplest example is the case where $\mathbb{X}$ is determined by an unbiased coin flip. Specifically, consider what happens if $\mathbb{X}$ takes value $\mathbb{X}_1$ if the coin comes up heads, otherwise it takes value $\mathbb{X}_2$. Then by the law of total probability, $\hat{\beta}$ has density $\frac{f_1}{2} + \frac{f_2}{2}$ where f_1, f_2 are densities of Gaussians $\mathcal{N}\left(\beta, \sigma^2(\mathbb{X}_1^T\mathbb{X}_1)^{-1}\right), \mathcal{N}\left(\beta, \sigma^2(\mathbb{X}_2^T\mathbb{X}_2)^{-1}\right)$ respectively. If $\mathbb{X}_1 \neq \mathbb{X}_2$, this is not a Gaussian distribution, but a "mixture" of two Gaussians. (Note: this is not to be confused with the density of the **sum of two Gaussian random variables**, which IS a Gaussian random variable. Summing the densities is different from summing the random variables!) In general, it can be very difficult to write down the distribution of $\hat{\beta}$ in terms of the distribution of $\mathbb{X}$.

Submit You have used 1 of 1 attempt

Answers are displayed within the problem

Discussion

Show Discussion

Topic: Unit 6 Linear Regression:Lecture 20: Linear Regression 2 / 7. Distribution of the Least Square Estimator