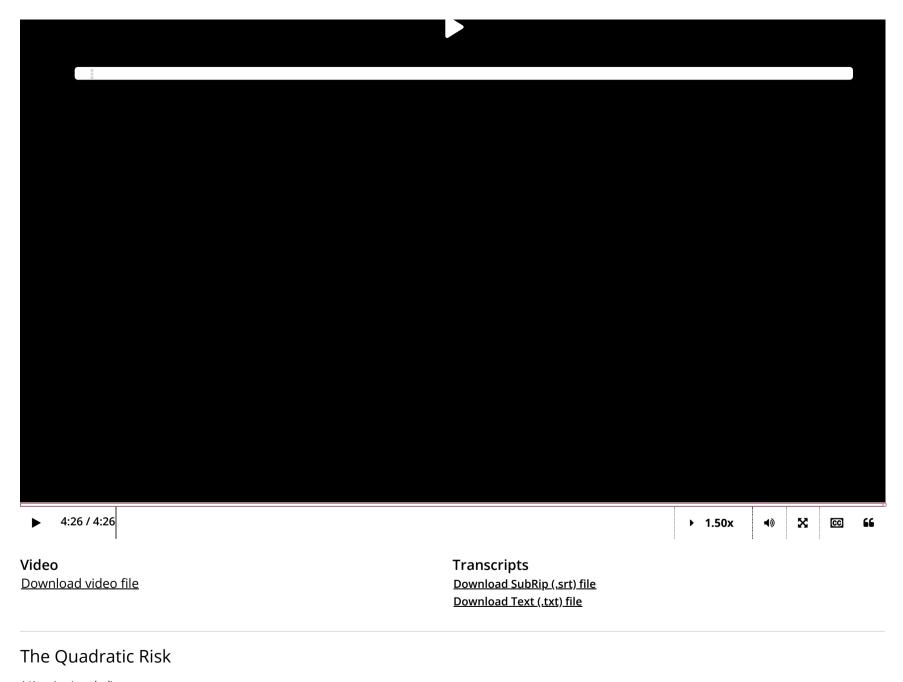


<u>Course</u> > <u>Unit 6 Linear Regression</u> > <u>Lecture 20: Linear Regression 2</u> > 9. Quadratic Risk and Variance

9. Quadratic Risk and Variance **Quadratic Risk**





1/1 point (graded)

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

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

What happens to these errors as σ^2 increases?

- igcup The error in $\hat{m{eta}}$ increases, but the prediction error decreases.
- igcap The error in $\hat{oldsymbol{eta}}$ decreases, but the prediction error increases.
- Both errors decrease.
- Both errors increase.



Solution:

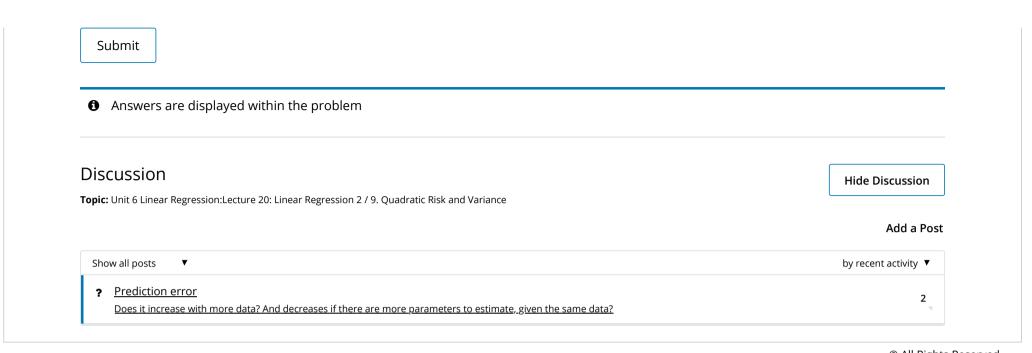
 σ^2 is the variance of each coordinate of ϵ . As the σ^2 increases, the data becomes more noisy. In particular, the task of estimating $\hat{\beta}$ ought to become harder, and it is intuitive that \mathbf{Y} becomes further from the prediction $\mathbb{X}\hat{\beta}$. To make this concrete, recall the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

$$\left\|\mathbb{E}[\|\hat{oldsymbol{eta}}-oldsymbol{eta}
ight\|_2^2=\sigma^2\mathrm{tr}\left(\left(\mathbb{X}^T\mathbb{X}
ight)^{-1}
ight)$$

$$\mathbb{E}\left[\left\|\mathbf{Y}-\mathbb{X}\hat{oldsymbol{eta}}
ight\|_{2}^{2}
ight]=\sigma^{2}\left(n-p
ight)$$

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

You have used 1 of 3 attempts



© All Rights Reserved