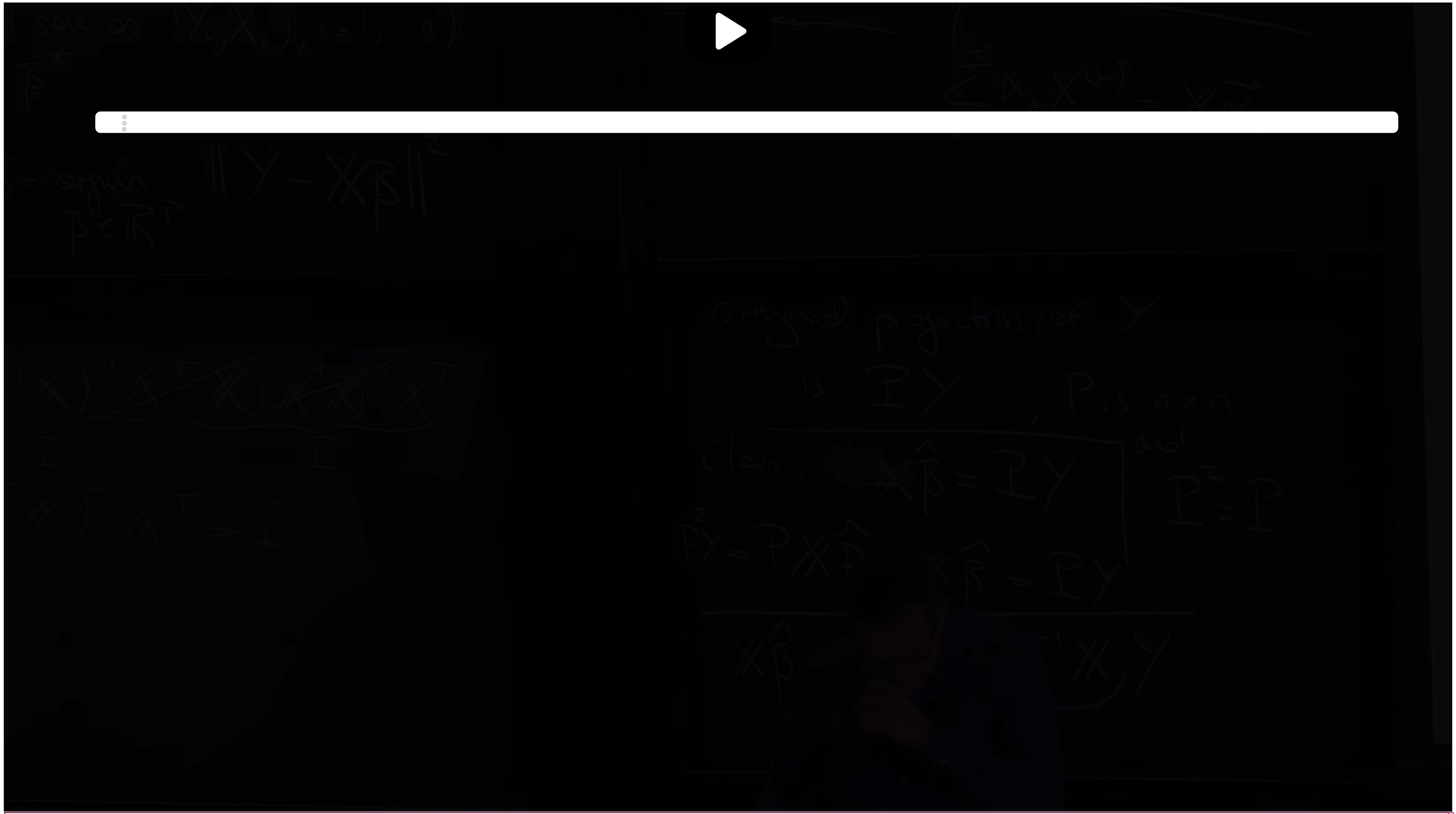
4. Geometric Interpretation of Linear Regression

Geometric Interpretation of Linear Regression



11:46 / 11:46 1.50x

Video

[Download video file]

Transcripts

[Download SubRip (.srt) file]

[Download Text (.txt) file]

Note: Also see recitation 13 *Hypothesis testing in linear regression* for some explanation and comments on the linear algebra. You will not be tested on the geometric interpretation on multivariate linear regression, ie. when $\beta \in \mathbb{R}^p$ for $p > 1$.

Geometric Interpretation of Linear Regression

1/1 point (graded)

Let $\mathrm{rank}\,(\mathbb{X}) = p$, so that $\mathbb{X}^T\mathbb{X}$ is invertible and the LSE $\hat{\beta}$ uniquely exists. The statistical interpretation here is that the product $\mathbb{X}\hat{\beta}$ provides the "best" *prediction* $\hat{\mathbf{Y}}$, in the sense that it minimizes the squared error $\|\mathbf{Y} - \hat{\mathbf{Y}}\|_2^2$. It also comes with a natural geometric interpretation, which we will now demonstrate.

Recall that the formula for the LSE is $\hat{\beta} = (\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T\mathbf{Y}$. Plugging this in gives

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

Thus, the prediction $\hat{\mathbf{Y}}$ is some linear transformation of the observed values $\mathbf{Y}$ via some matrix $\mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T$. Let $P = \mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T$. Which of the following statements about P, $\mathbb{X}$, $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ are true?

- [ ] $\mathbf{Y}$ is a linear combination of the columns of $\mathbb{X}$.
- [x] $\hat{\mathbf{Y}}$ is a linear combination of the columns of $\mathbb{X}$.
- [x] $P^2 = P$.
- [x] $P\hat{\mathbf{Y}} = \hat{\mathbf{Y}}$.

✔

Solution:

All choices are true statements, except for "$\mathbf{Y}$ **is in the column space of** $\mathbb{X}$".

The distinction between $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ being in the column space of $\mathbb{X}$ demonstrates that the data points give $\mathbf{Y}$ that may not be perfectly on the subspace generated by the columns of $\mathbb{X}$. However, the predictions $\hat{\mathbf{Y}}$ ought to be in the column space of $\mathbb{X}$, since predictions look like $\hat{Y} = X^T\beta$, which are **linear functions** of $\mathbf{X}$.

First, observe that since the left-most side in the product $\mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T$ is $\mathbb{X}$, so $\hat{\mathbf{Y}}$ is in the column space of $\mathbb{X}$. There is no such restriction on $\mathbf{Y}$, as it is being multiplied by P regardless of whether it lies in the subspace spanned by the columns of $\mathbb{X}$.
A direct calculation reveals

$$\begin{aligned} P^2 &= P \cdot P \\ &= (\mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T)\,(\mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T) \\ &= \mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}\,(\mathbb{X}^T\mathbb{X})\,(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T \\ &= \mathbb{X}(\mathbb{X}^T\mathbb{X})^{-1}\mathbb{X}^T = P. \end{aligned}$$

Such a matrix is commonly called a **projection** (or **idempotent**) matrix. This gives us the geometric interpretation of linear regression: **$\hat{\mathbf{Y}}$ is the orthogonal projection of $\mathbf{Y}$ onto the column space of $\mathbb{X}$.**
Finally, observe that $P\hat{\mathbf{Y}} = P\,(P\mathbf{Y}) = P\mathbf{Y} = \hat{\mathbf{Y}}$. This is natural: if we apply a projection to a vector, then apply the projection again, it should be the same as if we had applied it only once.

Submit

You have used 1 of 3 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 6 Linear Regression:Lecture 20: Linear Regression 2 / 4. Geometric Interpretation of Linear Regression

Add a Post

Show all posts ▼ by recent activity ▼

I'm still learning linear algebra, so this may be a stupid question, but.. 3