Stat 206: Linear Models

Lecture 7

October 16, 2019

Recap: Simple Linear Regression in Matrix Form

$$\underset{n\times 1}{\mathbf{Y}} = \underset{n\times 2}{\mathbf{X}}\ \underset{2\times 1}{\boldsymbol{\beta}} + \underset{n\times 1}{\boldsymbol{\epsilon}}.$$

- $\mathbf{E}\{\boldsymbol{\epsilon}\} = \mathbf{0}_n,\ \boldsymbol{\sigma}^2\{\boldsymbol{\epsilon}\} = \sigma^2\mathbf{I}_n.$
- Normal error model: $\boldsymbol{\epsilon} \sim \mathrm{Normal}_n(\mathbf{0}_n, \sigma^2\mathbf{I}_n)$.
- **LS estimators**:

$$\hat{\boldsymbol{\beta}} = \begin{bmatrix}\hat{\beta}_0\\ \hat{\beta}_1\end{bmatrix} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}. \tag{1}$$

- Fitted values and residuals:

$$\widehat{\mathbf{Y}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \mathbf{H}\mathbf{Y}, \quad \mathbf{e} = (\mathbf{I}_n - \mathbf{H})\mathbf{Y}.$$

- Hat matrix: $\mathbf{H} := \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, is a projection matrix.

Recap: Column Space of the Design Matrix X

- The design matrix
$$\mathbf{X} = (\mathbf{1}_n, \mathbf{x}).$$
- $\langle X \rangle = \{c_0 \mathbf{1}_n + c_1 \mathbf{x} = \mathbf{X}\mathbf{c} : c_0, c_1 \in R, \mathbf{c} = (c_0, c_1)^T\}$, is the linear subspace of $\mathbf{R}^n$ generated by the columns of $\mathbf{X}$.

Geometric Interpretation of Linear Regression

The hat matrix $\mathbf{H}$ projects a vector in $\mathbf{R}^n$ to the column space $\langle X \rangle$ of the design matrix $\mathbf{X}$: for any $\mathbf{w} \in \mathbf{R}^n$

- $\mathbf{Hw} \in \langle X \rangle$, i.e., there exists $c_0, c_1 \in \mathbf{R}$ such that $\mathbf{Hw} = c_0 \mathbf{1}_n + c_1 \mathbf{x}$.
- $\mathbf{w} - \mathbf{Hw} \perp \langle X \rangle$, i.e., for any $\mathbf{v} \in \langle X \rangle$, the inner product $< \mathbf{w} - \mathbf{Hw}, \mathbf{v} > = (\mathbf{w} - \mathbf{Hw})^T \mathbf{v} = 0$.

- $\widehat{\mathbf{Y}} = \mathbf{HY}$: the fitted values vector is the column space of $\mathbf{X}$:

- $\mathbf{e} = \mathbf{Y} - \mathbf{HY}$: the residuals vector is the column space of $\mathbf{X}$.
- So

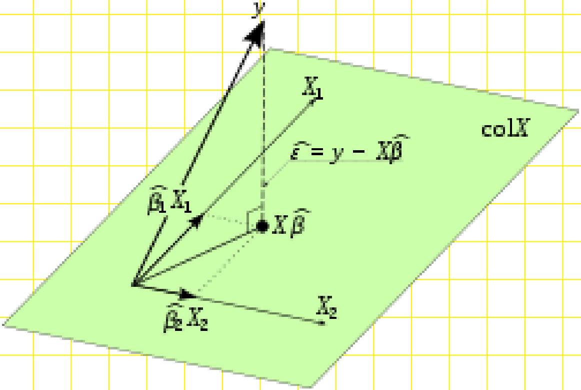
- $\widehat{\mathbf{Y}} = \mathbf{H}\mathbf{Y} \in \langle X \rangle$: the fitted values vector is in the column space of $\mathbf{X}$:

$$\widehat{\mathbf{Y}} = \hat{\beta}_0 \mathbf{1}_n + \hat{\beta}_1 \mathbf{x}.$$

- $\mathbf{e} = \mathbf{Y} - \mathbf{H}\mathbf{Y} \perp \langle X \rangle$: the residuals vector is orthogonal to the column space of $\mathbf{X}$.
- Since $\mathbf{1}_n, \mathbf{x}, \widehat{\mathbf{Y}} \in \langle X \rangle$, so

$$\begin{aligned}
< \mathbf{e}, \mathbf{1}_n > &= \sum_{i=1}^{n} e_i = 0 \\
< \mathbf{e}, \mathbf{x} > &= \sum_{i=1}^{n} X_i e_i = 0 \\
< \mathbf{e}, \widehat{\mathbf{Y}} > &= \sum_{i=1}^{n} \hat{Y}_i e_i = 0
\end{aligned}$$

Figure: Orthogonal projection of response vector **Y** onto the linear subspace of $\mathbb{R}^n$ generated by the columns of the design matrix **X**.



LS Estimators: Expectations

- LS estimators are unbiased estimators :

- Expectation of the fitted values:

- Expectation of the residuals:

LS Estimators: Expectations

- LS estimators are unbiased estimators :

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

- Expectation of the fitted values:

$$\mathbf{E}\{\widehat{\mathbf{Y}}\} = \mathbf{E}\{\mathbf{X}\hat{\beta}\} = \mathbf{X}\mathbf{E}\{\hat{\beta}\} = \mathbf{X}\beta = \mathbf{E}\{\mathbf{Y}\}.$$

- Expectation of the residuals:

$$\mathbf{E}\{\mathbf{e}\} = \mathbf{E}\{\mathbf{Y} - \widehat{\mathbf{Y}}\} = \mathbf{E}\{\mathbf{Y}\} - \mathbf{E}\{\widehat{\mathbf{Y}}\} = \mathbf{0}_n.$$

LS Estimators: Variance-covariance Matrices

- Variance-covariance of the LS estimators:

  What is the covariance between $\hat{\beta}_0$ and $\hat{\beta}_1$? What happens if $\overline{X} = 0$?

- Variance-covariance of fitted values:

- Variance-covariance of residuals:

  Are residuals uncorrelated? Do they have the same variance?

LS Estimators: Variance-covariance Matrices

- Variance-covariance of the LS estimators:

$$\begin{aligned}
\boldsymbol{\sigma^2}\{\hat{\boldsymbol{\beta}}\} &= \boldsymbol{\sigma^2}\{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}\} = \left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right)\boldsymbol{\sigma^2}\{\mathbf{Y}\}\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right)' \\
&= \sigma^2(\mathbf{X}'\mathbf{X})^{-1} = \sigma^2\begin{bmatrix} \frac{1}{n} + \frac{\overline{X}^2}{\sum_{i=1}^n (X_i-\overline{X})^2} & -\frac{\overline{X}}{\sum_{i=1}^n (X_i-\overline{X})^2} \\ -\frac{\overline{X}}{\sum_{i=1}^n (X_i-\overline{X})^2} & \frac{1}{\sum_{i=1}^n (X_i-\overline{X})^2} \end{bmatrix}
\end{aligned}$$

What is the covariance between $\hat{\beta}_0$ and $\hat{\beta}_1$? What happens if $\overline{X} = 0$?

- Variance-covariance of fitted values:

$$\boldsymbol{\sigma^2}\{\widehat{\mathbf{Y}}\} = \mathbf{H}\boldsymbol{\sigma^2}\{\mathbf{Y}\}\mathbf{H}' = \sigma^2\mathbf{H}.$$

- Variance-covariance of residuals:

$$\boldsymbol{\sigma^2}\{\mathbf{e}\} = (\mathbf{I}_n - \mathbf{H})\boldsymbol{\sigma^2}\{\mathbf{Y}\}(\mathbf{I}_n - \mathbf{H})' = \sigma^2(\mathbf{I}_n - \mathbf{H}).$$

Are residuals uncorrelated? Do they have the same variance?

Sum of Squares in Matrix Form

Error sum of squares:

$$SSE = \sum_{i=1}^{n} e_i^2.$$

- Matrix form:

- Recall that $\mathbf{I}_n - \mathbf{H}$ is a matrix.
- $df(SSE) =$

Sum of Squares in Matrix Form

Error sum of squares:

$$SSE = \sum_{i=1}^{n} e_i^2.$$

- Matrix form:

$$SSE = \mathbf{e}'\mathbf{e} = \mathbf{Y}'(\mathbf{I}_n - \mathbf{H})'(\mathbf{I}_n - \mathbf{H})\mathbf{Y} = \mathbf{Y}'(\mathbf{I}_n - \mathbf{H})\mathbf{Y}.$$

- Recall that $\mathbf{I}_n - \mathbf{H}$ is a projection matrix. *Which space it projects to?*
- $df(SSE) = rank(\mathbf{I}_n - \mathbf{H}) = n - 2.$