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 $\mathbf{X}$.

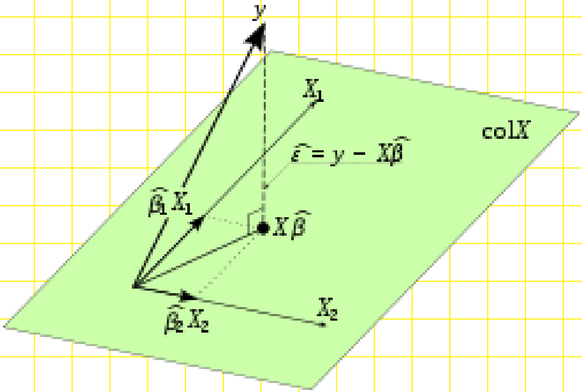
Geometric Interpretation of Linear Regression

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

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: 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?

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.
- $df(SSE) =$

Total sum of squares:

$$SSTO = \sum_{i=1}^{n}(Y_i - \overline{Y})^2 = \sum_{i=1}^{n} Y_i^2 - n(\overline{Y})^2.$$

- Matrix form:

- Note $\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n$ is a projection matrix:

$$\mathbf{J}_n = \mathbf{1}_n \mathbf{1}_n' = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}.$$

- $df(SSTO) =$

Regression sum of squares : $SSR = \sum_{i=1}^{n}(\widehat{Y}_i - \overline{Y})^2$.

- Matrix form: $\overline{\mathbf{Y}} =$

- Note $\mathbf{H} - \frac{1}{n}\mathbf{J}_n$ is a projection matrix:
- $df(SSR) =$

Recap: Sum of Squares in Matrix Form

- $$SSTO = \mathbf{Y}'\left(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n\right)\mathbf{Y}.$$
- $$SSE = \mathbf{Y}'(\mathbf{I}_n - \mathbf{H})\mathbf{Y}.$$
- $$SSR = \mathbf{Y}'(\mathbf{H} - \frac{1}{n}\mathbf{J}_n)\mathbf{Y}.$$

$E(SSE)$

Properties of Projection Matrices

Optional Reading material.

- They have eigen-decomposition of the form:

$$Q\Lambda Q^T,$$

  where Q is an orthogonal matrix of eigenvectors and Λ is a diagonal matrix of eigenvalues.
- Their eigenvalues are either 1 or 0.
- The number of nonzero eigenvalues equals to trace of the matrix equals to the rank.
- For simple linear regression:

$$rank(\mathbf{H}) = 2, \quad rank(\mathbf{I}_n - \mathbf{H}) = n - 2.$$

Sampling Distribution of SSE

Optional Reading material (cont'd).

- $\mathbf{I}_n - \mathbf{H}$ is a projection matrix with rank $n - 2 \Longrightarrow$

$$\mathbf{I}_n - \mathbf{H} = \mathbf{Q}^T \mathbf{\Lambda} \mathbf{Q},$$

  where $\mathbf{\Lambda} = \text{diag}\{1, \cdots, 1, 0, 0\}$ and $\mathbf{Q}$ is an orthogonal matrix.

- $(\mathbf{I}_n - \mathbf{H})\mathbf{X} = \mathbf{0} \Longrightarrow$

$$\mathbf{e} = (\mathbf{I}_n - \mathbf{H})\mathbf{Y} = (\mathbf{I}_n - \mathbf{H})(\mathbf{X}\beta + \epsilon) = (\mathbf{I}_n - \mathbf{H})\epsilon$$

Optional Reading material (cont'd).

- $SSE = \mathbf{e}^T\mathbf{e} = \boldsymbol{\epsilon}^T(\mathbf{I}_n - \mathbf{H})\boldsymbol{\epsilon} = (\mathbf{Q}\boldsymbol{\epsilon})^T\boldsymbol{\Lambda}(\mathbf{Q}\boldsymbol{\epsilon})$.
- Let $\mathbf{z} = \mathbf{Q}\boldsymbol{\epsilon}$, then

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

- Moreover

$$\mathbf{E}(\mathbf{z}) = \mathbf{Q}\mathbf{E}\{\boldsymbol{\epsilon}\} = \mathbf{0}, \quad \boldsymbol{\sigma}^2\{\mathbf{z}\} = \mathbf{Q}\boldsymbol{\sigma}^2\{\boldsymbol{\epsilon}\}\mathbf{Q}^T = \sigma^2\mathbf{Q}\mathbf{Q}^T = \sigma^2\mathbf{I}_n.$$

  So under Normal error model, z_is are i.i.d. $N(0, \sigma^2)$.
- So $SSE \sim \sigma^2\chi^2_{(n-2)}$.