

STAT 3690 Lecture 16

zhiyanggeezhou.github.io

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca)

Mar 07, 2022

What is a linear model?

- Responses are linear functions with respect to unknown parameters.

Univariate/multiple linear regression

- Interested in the relationship between random scalar Y and random q -vector $[X_1, \dots, X_q]^\top$
- Model
 - Population version: $Y \mid X_1, \dots, X_q \sim ([1, X_1, \dots, X_q]\boldsymbol{\beta}, \sigma^2)$, where $\boldsymbol{\beta} = [\beta_0, \dots, \beta_q]^\top$, i.e.,
 - * $E(Y \mid X_1, \dots, X_q) = [1, X_1, \dots, X_q]\boldsymbol{\beta} = \beta_0 + \sum_{j=1}^q X_j\beta_j$
 - * $\text{var}(Y \mid X_1, \dots, X_q) = \sigma^2$
 - Sample version $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$
 - * $\mathbf{Y} = [Y_1, \dots, Y_n]^\top$ and design matrix

$$\mathbf{X} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{q1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \cdots & X_{nq} \end{bmatrix}_{n \times (q+1)}$$

- Independent realizations $[Y_i, X_{i1}, \dots, X_{iq}]^\top \sim [Y, X_1, \dots, X_q]^\top, i = 1, \dots, n$
- $\text{rk}(\mathbf{X}) = q + 1 < p + q + 1 \leq n$
- * $\boldsymbol{\varepsilon} = [\varepsilon_1, \dots, \varepsilon_n]^\top \sim (\mathbf{0}_n, \sigma^2 \mathbf{I}_n)$

-
- Least squares (LS) estimation (no need of normality)
 - $\hat{\boldsymbol{\beta}}_{\text{LS}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y}$
 - $\hat{\sigma}_{\text{LS}}^2 = (n - q - 1)^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})^\top (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = (n - q - 1)^{-1} \mathbf{Y}^\top (\mathbf{I} - \mathbf{H}) \mathbf{Y}$
 - * Hat matrix $\mathbf{H} = [h_{ij}]_{n \times n} = \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top$
 - Symmetric
 - Idempotent: $\mathbf{H}^2 = \mathbf{H}\mathbf{H} = \mathbf{H}$
 - $\text{rk}(\mathbf{H}) = \text{rk}(\mathbf{X})$
 - Each eigenvalue of \mathbf{H} is either zero or one

-
- Maximum likelihood (ML) estimation (in need of normality)
 - $\hat{\boldsymbol{\beta}}_{\text{ML}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y} = \hat{\boldsymbol{\beta}}_{\text{LS}}$
 - $\hat{\sigma}_{\text{ML}}^2 = n^{-1} \mathbf{Y}^\top (\mathbf{I} - \mathbf{H}) \mathbf{Y} = n^{-1} (n - q - 1) \hat{\sigma}_{\text{LS}}^2$

-
- Inference (in need of normality)

- 100(1 – α)% confidence interval for $\mathbf{a}^\top \boldsymbol{\beta}$ with known $\mathbf{a} \in \mathbb{R}^{q+1}$:

$$\mathbf{a}^\top \hat{\boldsymbol{\beta}}_{\text{ML}} \pm t_{1-\alpha/2, n-q-1} \hat{\sigma}_{\text{ML}} [\mathbf{a}^\top (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{a}]^{1/2}$$

- 100(1 – α)% prediction interval for $Y_0 = \mathbf{X}_0 \boldsymbol{\beta} + \varepsilon_0$, given $\mathbf{X}_0 = [1, X_{01}, \dots, X_{0q}] \in \mathbb{R}^{q+1}$:

$$\mathbf{X}_0 \hat{\boldsymbol{\beta}}_{\text{ML}} \pm t_{1-\alpha/2, n-q-1} \hat{\sigma}_{\text{ML}} [1 + \mathbf{X}_0 (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}_0^\top]^{1/2}$$