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 $[X_1, \ldots, X_q]^\top$
- Model
  - Population version: $Y \mid X_1, \ldots, X_q \sim ([1, X_1, \ldots, X_q]\boldsymbol{\beta}, \sigma^2)$, where $\boldsymbol{\beta} = [\beta_0, \ldots, \beta_q]^\top$, i.e.,
    * $\mathrm{E}(Y \mid X_1, \ldots, X_q) = [1, X_1, \ldots, X_q]\boldsymbol{\beta} = \beta_0 + \sum_{j=1}^q X_j\beta_j$
    * $\mathrm{var}(Y \mid X_1, \ldots, X_q) = \sigma^2$
  - Sample version $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$
    * $\mathbf{Y} = [Y_1, \ldots, Y_n]^\top$ 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 $[Y_i, X_{i1}, \ldots, X_{iq}]^\top \sim [Y, X_1, \ldots, X_q]^\top$, $i = 1, \ldots, n$
      · $\mathrm{rk}(\mathbf{X}) = q + 1 < p + q + 1 \le n$
    * $\boldsymbol{\varepsilon} = [\varepsilon_1, \ldots, \varepsilon_n]^\top \sim (\mathbf{0}_n, \sigma^2\mathbf{I}_n)$

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

---

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

- Inference (in need of normality)

- New realization $[Y_0, X_{01}, \dots, X_{0q}]^\top \sim [Y, X_1, \dots, X_q]^\top$
- New design matrix $\mathbf{X}_0 = [1, X_{01}, \dots, X_{0q}]^\top$
- $100(1-\alpha)\%$ confidence interval for $\mathrm{E}(Y_0|X_{01}, \dots, X_{0q}) = \mathbf{X}_0^\top \boldsymbol{\beta}$:

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

- $100(1-\alpha)\%$ prediction interval for $Y_0 = \mathbf{X}_0^\top \boldsymbol{\beta} + \varepsilon_0$:

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