STAT 3690 Lecture 17

zhiyanggeezhou.github.io

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca)

Mar 09, 2022

Multivariate linear regression

- Interested in the relationship between random q-vector $[Y_1, \ldots, Y_p]^\top$ and random q-vector $[X_1, \ldots, X_q]^\top$
- Model
  - Population version: $[Y_1, \ldots, Y_p]^\top \mid X_1, \ldots, X_q \sim (\mathbf{B}^\top [1, X_1, \ldots, X_q]^\top, \sigma^2)$, where $\mathbf{B} = [\beta_{kj}]_{(q+1)\times p}$, i.e.,
    * $\mathrm{E}([Y_1, \ldots, Y_p]^\top \mid X_1, \ldots, X_q) = \mathbf{B}^\top [1, X_1, \ldots, X_q]^\top$
    * $\mathrm{cov}([Y_1, \ldots, Y_p]^\top \mid X_1, \ldots, X_q) = \mathbf{\Sigma} > 0$, i.e., the conditional covariance of $[Y_1, \ldots, Y_p]^\top$ does not depend on $X_1, \ldots, X_q$
  - Sample version
$$\underset{n\times p}{\mathbf{Y}} = \underset{n\times(q+1)}{\mathbf{X}}\ \underset{(q+1)\times p}{\mathbf{B}} + \underset{n\times p}{\mathbf{E}}$$
    * $\mathbf{Y} = [Y_{ij}]_{n\times p}$
    * Design matrix
$$\mathbf{X} = \begin{bmatrix} 1 & X_{11} & \cdots & X_{q1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{n1} & \cdots & X_{nq} \end{bmatrix}_{n\times(q+1)}$$
      · $\mathrm{rk}(\mathbf{X}) = q + 1 < p + q + 1 \le n$
    * $\mathbf{E} = [\mathbf{E}_{1\cdot}, \ldots, \mathbf{E}_{n\cdot}]^\top$, where $\mathbf{E}_{i\cdot}^\top$ is the ith row of $\mathbf{E}$
    * Assume the independence across i, i.e.,
      · $[Y_{i1}, \ldots, Y_{ip}, X_{i1}, \ldots, X_{iq}]^\top \overset{\text{iid}}{\sim} [Y_1, \ldots Y_p, X_1, \ldots, X_q]^\top$
      · $\mathbf{E}_{i\cdot} \overset{\text{iid}}{\sim} (\mathbf{0}_p, \mathbf{\Sigma})$

- Relationship with univariate linear regression
  - If $\mathbf{\Sigma}$ is diagonal, the multivariate model reduces to $\mathbf{Y}_{\cdot j} = \mathbf{X}\mathbf{B}_{\cdot j} + \mathbf{E}_{\cdot j}$, $j = 1, \ldots, p$
    * $\mathbf{Y}_{\cdot j}$: the jth column of $\mathbf{Y}$
    * $\mathbf{B}_{\cdot j}$: the jth column of $\mathbf{B}$
    * $\mathbf{E}_{\cdot j} \sim (\mathbf{0}_n, \sigma_{jj}^2 \mathbf{I}_n)$
      · σ_{jj}^2: (j, j)-entry of $\mathbf{\Sigma}$

- Relationship with MANOVA
  - MANOVA models can be expressed as multivariate linear regression with carefully selected dummy (explanatory) variables.

Exercise: translate the following 1-way MANOVA model

$$\mathbf{Y}_{ij} = \boldsymbol{\mu} + \boldsymbol{\tau}_i + \mathbf{E}_{ij}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, n_i$$

into a multivariate linear regression model, where $\mathbf{E}_{ij} \overset{\text{iid}}{\sim} MVN_p(\mathbf{0}, \boldsymbol{\Sigma})$ and $\sum_i \boldsymbol{\tau}_i = 0$.

- Least squares (LS) estimation (no need of (conditional) normality)
  - $\hat{\mathbf{B}}_{\text{LS}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y}$
    - $\text{E}(\hat{\mathbf{B}}_{\text{LS}}) = \mathbf{B}$
  - $\hat{\boldsymbol{\Sigma}}_{\text{LS}} = (n - q - 1)^{-1} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})^\top (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = (n - q - 1)^{-1} \mathbf{Y}^\top (\mathbf{I} - \mathbf{H}) \mathbf{Y}$
    - $\text{E}(\hat{\boldsymbol{\Sigma}}_{\text{LS}}) = \boldsymbol{\Sigma}$

- Maximum likelihood (ML) estimation (in need of (conditional) normality)
  - $\hat{\mathbf{B}}_{\text{ML}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y} = \hat{\mathbf{B}}_{\text{LS}}$
  - $\hat{\boldsymbol{\Sigma}}_{\text{ML}} = n^{-1} \mathbf{Y}(\mathbf{I} - \mathbf{H}) \mathbf{Y} = n^{-1}(n - q - 1) \hat{\boldsymbol{\Sigma}}_{\text{LS}}$
    - Given $\mathbf{X}$, $n\hat{\boldsymbol{\Sigma}}_{\text{ML}} \sim W_p(\boldsymbol{\Sigma}, n - q - 1)$

- Inference (in need of (conditional) normality)
  - Inference on $\mathbf{B}^\top \boldsymbol{a}$, given $\boldsymbol{a} \in \mathbb{R}^{q+1}$
    - Estimator $\hat{\mathbf{B}}_{\text{ML}}^\top \boldsymbol{a}$
    - $100(1 - \alpha)\%$ confidence region for $\mathbf{B}^\top \boldsymbol{a}$

$$\left\{ \boldsymbol{u} \in \mathbb{R}^p : (\boldsymbol{u} - \widehat{\mathbf{B}}_{\text{ML}}^\top \boldsymbol{a})^\top \widehat{\boldsymbol{\Sigma}}_{\text{LS}}^{-1} (\boldsymbol{u} - \hat{\mathbf{B}}_{\text{ML}}^\top \boldsymbol{a}) \leq \frac{p(n - q - 1)}{n - q - p} \boldsymbol{a}^\top (\mathbf{X}^\top \mathbf{X})^{-1} \boldsymbol{a} F_{1-\alpha, p, n-p-q} \right\}$$

  - Inference on $\mathbf{Y}_0 = \mathbf{B}^\top \mathbf{X}_0 + \mathbf{E}_0$ with a new observation vector $\mathbf{X}_0 = [1, X_{01}, \ldots, X_{0q}]^\top \in \mathbb{R}^{q+1}$
    - Prediction $\hat{\mathbf{Y}}_0 = \mathbf{B}_{\text{ML}}^\top \mathbf{X}_0$
    - $100(1 - \alpha)\%$ prediction region for $\mathbf{Y}_0$

$$\left\{ \boldsymbol{u} \in \mathbb{R}^p : (\boldsymbol{u} - \hat{\mathbf{Y}}_0)^\top \widehat{\boldsymbol{\Sigma}}_{\text{LS}}^{-1} (\boldsymbol{u} - \hat{\mathbf{Y}}_0) \leq \frac{p(n - q - 1)}{n - q - p} \{1 + \mathbf{X}_0^\top (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}_0\} F_{1-\alpha, p, n-p-q} \right\}$$

  - Inference on $\boldsymbol{a}^\top \mathbf{Y}_0 = \boldsymbol{a}^\top \mathbf{B}^\top \mathbf{X}_0 + \boldsymbol{a}^\top \mathbf{E}_0$, given $\boldsymbol{a} \in \mathbb{R}^p$ and a new observation vector $\mathbf{X}_0 = [1, X_{01}, \ldots, X_{0q}]^\top \in \mathbb{R}^{q+1}$
    - Prediction $\boldsymbol{a}^\top \hat{\mathbf{Y}}_0 = \boldsymbol{a}^\top \mathbf{B}_{\text{ML}}^\top \mathbf{X}_0$
    - $100(1 - \alpha)\%$ Scheffé's simultaneous prediction interval for $\boldsymbol{a}^\top \mathbf{Y}_0$

$$\boldsymbol{a}^\top \hat{\mathbf{Y}}_0 \pm \sqrt{\boldsymbol{a}^\top \hat{\boldsymbol{\Sigma}}_{\text{LS}} \boldsymbol{a} \frac{p(n - q - 1)}{n - q - p} \{1 + \mathbf{X}_0^\top (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}_0\} F_{1-\alpha, p, n-p-q}}$$