9 Generalized Least Squares

What happens if we relax the assumption that $\text{cov}(\mathbf{Y}) = \sigma^2\mathbf{I}$?

9.1 Example: (Clustered data). Let

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \mathbf{Y}_2 \\ \vdots \\ \mathbf{Y}_K \end{pmatrix},$$

where $\mathbf{Y}_i = (Y_{i1}, \ldots, Y_{in_i})'$ is a vector of responses on the ith cluster (patient, household, school, etc). Assuming clusters are independent,

$$\text{cov}(\mathbf{Y}) = \begin{pmatrix} \mathbf{V}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{V}_K \end{pmatrix},$$

where we might assume a common variance σ^2 and common pairwise correlation ρ within a cluster, i.e. an exchangeable correlation structure:

$$cov(\mathbf{Y}_i) = \sigma^2\mathbf{V}_i = \sigma^2 \begin{pmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \rho \\ \rho & \cdots & \rho & 1 \end{pmatrix}_{n_i \times n_i}$$

In general, let $\text{cov}(\mathbf{Y}) = \sigma^2\mathbf{V}$, $\mathbf{V}$is known. In practice, we will also have to estimate $\mathbf{V}$ (e.g. the correlation parameter ρ in the exchangeable case).

9.2 Theorem: Let $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where rank $(\mathbf{X}_{n\times p}) = p$, $E[\boldsymbol{\varepsilon}] = \mathbf{0}$, $\text{cov}(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{V}$, with known p.d. $\mathbf{V}$. There exists a transformation of $\mathbf{Y}$ to a new response vector which has covariance matrix $\sigma^2\mathbf{I}$. Least squares applied to the transformed $\mathbf{Y}$ yields

$$\boldsymbol{\beta}^* = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y},$$

the Generalized Least Squares (GLS) estimate.

9.3 Theorem: Properties of $\boldsymbol{\beta}^*$:

(a) $E[\boldsymbol{\beta}^*] = \boldsymbol{\beta}$,

(b) $\text{cov}(\boldsymbol{\beta}^*) = \sigma^2(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$,

(c) $RSS = (\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}^*)'\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta}^*)$.

Let $\boldsymbol{\beta}^* = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Y}$ be the generalized least squares (GLS) estimate, and $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$ be the ordinary least squares (OLS) estimate.

9.4 Theorem: Under the conditions of Theorem 9.2, the OLS estimate has the following properties:

(a) $E[\hat{\boldsymbol{\beta}}] = \boldsymbol{\beta}$,

(b) $\text{cov}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{V}\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}$.

9.5 Theorem: (Optimality of GLS estimates). If $E[Y] = \mathbf{X}\boldsymbol{\beta}$ and $\text{cov}(\mathbf{Y}) = \sigma^2\mathbf{V}$, then for any constant vector $\mathbf{a}$, $\mathbf{a}'\boldsymbol{\beta}^*$ is the BLUE of $\mathbf{a}'\boldsymbol{\beta}$.

9.6 Example: (Weighted least squares). Let $Y_1, \ldots, Y_n$ be independent, $E[Y_i] = \beta x_i$, and $\text{var}(Y_i) = \sigma^2 w_i^{-1}$. The GLS estimate of β is

$$\boldsymbol{\beta}^* = \frac{\sum_{i=1}^n w_i x_i Y_i}{\sum_{i=1}^n w_i x_i^2}.$$

The OLS estimate is

$$\hat{\boldsymbol{\beta}} = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}.$$

The variances are

$$\text{var}(\boldsymbol{\beta}^*) = \frac{\sigma^2}{\sum_{i=1}^n w_i x_i^2} \qquad \text{and} \qquad \text{var}(\hat{\boldsymbol{\beta}}) = \frac{\sigma^2 \sum_{i=1}^n \frac{x_i^2}{w_i}}{(\sum_{i=1}^n x_i^2)^2}.$$

9.7 Theorem: The GLS estimate and the OLS estimate are equal only when either one of the following conditions holds:

1. $\mathcal{R}(\mathbf{V}^{-1}\mathbf{X}) = \mathcal{R}(\mathbf{X})$.
2. $\mathcal{R}(\mathbf{V}\mathbf{X}) = \mathcal{R}(\mathbf{X})$.