9 Generalized Least Squares

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

9.1 Example: (Clustered data). Let

\[
Y = \begin{pmatrix}
Y_1 \\
Y_2 \\
\vdots \\
Y_K
\end{pmatrix},
\]

where \( Y_i = (Y_{i1}, \ldots, Y_{in_i})' \) is a vector of responses on the \( i \)th cluster (patient, household, school, etc). Assuming clusters are independent,

\[
\text{cov}(Y) = \begin{pmatrix}
V_1 & 0 & \cdots & 0 \\
0 & V_2 & \cdots & \vdots \\
\vdots & \vdots & \ddots & 0 \\
0 & \cdots & 0 & V_K
\end{pmatrix},
\]

where we might assume a common variance \( \sigma^2 \) and common pairwise correlation \( \rho \) within a cluster, i.e. an exchangeable correlation structure:

\[
\text{cov}(Y_i) = \sigma^2 V_i = \sigma^2 \begin{pmatrix}
1 & \rho & \cdots & \rho \\
\rho & 1 & \cdots & \vdots \\
\vdots & \vdots & \ddots & \rho \\
\rho & \cdots & \rho & 1
\end{pmatrix}_{n_i \times n_i}
\]

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

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

\[
\beta^* = (X'V^{-1}X)^{-1}X'V^{-1}Y,
\]

the Generalized Least Squares (GLS) estimate.

9.3 Theorem: Properties of \( \beta^* \):

(a) \( E[\beta^*] = \beta \),

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

(c) \( \text{RSS} = (Y - X\beta^*)'V^{-1}(Y - X\beta^*) \).
Let $\beta^* = (X'V^{-1}X)^{-1}X'V^{-1}Y$ be the generalized least squares (GLS) estimate, and $\hat{\beta} = (X'X)^{-1}X'Y$ 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) $E[\hat{\beta}] = \beta$,
(b) $\text{cov}(\hat{\beta}) = \sigma^2(X'X)^{-1}(X'VX)(X'X)^{-1}$.

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

**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 $\beta$ is

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

The OLS estimate is

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

The variances are

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

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

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