Longitudianl Data Analysis

Weighted least-squares (WLS) estimation

(no distribution assumption on $\mathbf{y}$) The **weighted least-squares** estimator of $\boldsymbol{\beta}$, using a symmetric *weight matrix*, W, is the value, $\tilde{\boldsymbol{\beta}}_W$, which minimize the quadratic form

$$(\mathbf{y} - X\boldsymbol{\beta})'W(\mathbf{y} - X\boldsymbol{\beta}).$$

Standard matrix manipulations give the explicit result

$$\tilde{\boldsymbol{\beta}}_W = (X'WX)^{-1}X'W\mathbf{y}$$

- it's an unbiased estimator of $\boldsymbol{\beta}$, whatever the choice of W.

- $Var(\tilde{\boldsymbol{\beta}}_W) = \sigma^2\{(X'WX)^{-1}X'W\}V\{WX(X'WX)^{-1}\}$.

1. If $W = I$, it reduces to the OLS estimator

   $$\tilde{\boldsymbol{\beta}}_I = (X'X)^{-1}X\mathbf{y},$$

   with

   $$Var(\tilde{\boldsymbol{\beta}}_I) = \sigma^2(X'X)^{-1}X'VX(X'X)^{-1}.$$

2. If $W = V^{-1}$, the estimator becomes the *MLE* (under the assumption of normal dist.), i.e.

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

   with

   $$Var(\hat{\boldsymbol{\beta}}) = \sigma^2(X'V^{-1}X)^{-1}.$$

- According to the [G-M Theorem] (https://en.wikipedia.org/wiki/Gauss–Markov_theorem), the MLE is the most efficient linear estimator for $\boldsymbol{\beta}$. However, to identify this optimal weighting matrix we need to know the complete correlation structure of the data – we don't need to know σ^2, because $\tilde{\boldsymbol{\beta}}_W$ is unchanged by proportional changes in all the elements of W.

- Also, because the correlation structure may be diffcult to identify in practice, it is of interest to ask how much loss of efficiency might result from using a different W.

- When we know the correlation structure is CS (uniform/exchangable), the OLS is fully efficient as the WLS estimator; an intuitive explanation is that with a common correlation between any two equally spaced measurements on the same unit, there is no reason to weight measurements differently.

Using OLS estimator is misleading when $V \neq I$

- In many circumstances where there is balanced design, the OLS estimator, $\tilde{\boldsymbol{\beta}}$, is perfectly satisfactory for point estimation. But this is not always the case. (example in book page 63.)

- Even when OLS is reasonably efficient, it is clear from the form of

  $$Var(\tilde{\boldsymbol{\beta}}_I) = \sigma^2 (X'X)^{-1} X'VX(X'X)^{-1}$$

  that interval estimation for $\boldsymbol{\beta}$ still requires information about $\sigma^2 V$, the variance matrix of the data. In particular, the usual formula for the variance of the least-squares estimator,

  $$Var(\tilde{\boldsymbol{\beta}}) = \sigma^2 (X'X)^{-1}$$

  assumes that $V = I$, the identity matrix, and can be seriously misleading when this is not so.

- A naive use of OLS would be to ignore the correlation structure in the data and to base interval estimation for $\boldsymbol{\beta}$ on the variance above with σ^2 replaced with its usual estimator, the residual mean square

  $$\tilde{\sigma}^2 = (nm - p)^{-1} (\mathbf{y} - X\tilde{\boldsymbol{\beta}})'(\mathbf{y} - X\tilde{\boldsymbol{\beta}}).$$

  There are two sources of error in this naive approach when $V \neq I$:

1. $Var(\tilde{\boldsymbol{\beta}})$ is wrong
2. $\tilde{\sigma}^2$ is no longer an unbiased esitmator of σ^2.

MLE under Gaussian assumptions

REML

(under Gaussian assumptions)