Generalized Least Squares (GLS) Estimation

Gaurav Rampuria (BS SDS, IIT Kanpur)

June 2, 2025

Introduction

Generalized Least Squares (GLS) is an extension of the Ordinary Least Squares (OLS) method for linear regression, designed to handle cases where the error terms are heteroskedastic (have non-constant variance) or correlated. GLS provides the Best Linear Unbiased Estimator (BLUE) when the usual OLS assumptions are violated.

Linear Regression Model

Consider the linear regression model:

$$y = X\beta + \varepsilon$$

where

- y is an $n \times 1$ vector of observations,
- X is an $n \times p$ matrix of regressors,
- β is a $p \times 1$ vector of unknown parameters,
- ε is an $n \times 1$ vector of errors.

OLS Assumptions:

- $\mathbb{E}[\varepsilon] = 0$
- $\mathrm{Cov}(\varepsilon) = \sigma^2 I$ (errors are uncorrelated and have equal variance)

GLS Context:

- $\mathbb{E}[\varepsilon] = 0$
- $\mathrm{Cov}(\varepsilon) = \Omega$ (errors may be correlated and/or have unequal variance; Ω is a known positive-definite matrix)

The GLS Estimator

GLS minimizes the weighted sum of squared residuals:

$$\hat{\beta}_{GLS} = \arg\min_{\beta} (y - X\beta)^\top \Omega^{-1} (y - X\beta)$$

The solution is:

$$\boxed{\hat{\beta}_{GLS} = (X^\top \Omega^{-1} X)^{-1} X^\top \Omega^{-1} y}$$

If $\Omega = \sigma^2 I$, this reduces to the OLS estimator.

Interpretation

GLS "whitens" the data by transforming the model so that the transformed errors are uncorrelated and have equal variance. This is often done by factoring $\Omega = CC^\top$ (e.g., Cholesky decomposition) and multiplying both sides by C^{-1}, so OLS can be applied to the transformed model.

Why Use GLS?

- **Efficiency:** When errors are heteroskedastic or correlated, OLS is unbiased but not efficient (not minimum variance). GLS is BLUE in these cases.
- **Correct Inference:** Standard errors and confidence intervals from OLS are invalid under heteroskedasticity or autocorrelation. GLS corrects this.

Special Cases

- **Weighted Least Squares (WLS):** If Ω is diagonal (errors uncorrelated but have different variances), GLS becomes WLS, where each observation is weighted by the inverse of its variance.
- **Feasible GLS (FGLS):** In practice, Ω is rarely known. FGLS estimates Ω from the data (often in two stages: estimate the model by OLS, estimate Ω from residuals, then re-estimate β by GLS using the estimated Ω).

Properties

- **Unbiasedness:** $\mathbb{E}[\hat{\beta}_{GLS}] = \beta$
- **Variance:** $\mathrm{Cov}[\hat{\beta}_{GLS}] = (X^\top \Omega^{-1} X)^{-1}$
- **BLUE:** Among all linear unbiased estimators, GLS has the minimum variance when Ω is correctly specified.

Summary Table

Method	Error Covariance	Estimator Formula	When to Use
OLS	$\sigma^2 I$	$(X^\top X)^{-1} X^\top y$	Errors are i.i.d.
GLS	Ω (known)	$(X^\top \Omega^{-1} X)^{-1} X^\top \Omega^{-1} y$	Errors are correlated / heteroskedastic, Ω known
FGLS	Ω (estimated)	Same as GLS but with estimated Ω	Errors are correlated / heteroskedastic, Ω unknown

References

- Rao, C.R. (1973). *Linear Statistical Inference and Its Applications.* Wiley, New York.
- StatLect: Generalized Least Squares. `https://www.statlect.com/fundamentals-of-statistics/generalized-least-squares`