Recap:

(i) Simple Model: $Y_i = \alpha + \beta X_i + u_i$

(ii) Multivariate Model: $Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i$

(iii) OLS Estimation Technique

(iv) Assumptions of the Model [$E(u_i) = 0$, $Var(u_i) = \sigma^2 \; \forall i$, ...].

(v) Properties of the OLS Estimates. [unbiaseness, variance].

(vi) Gauss Markov Theorem.

Under the classical assumptions OLS estimates are BLUE.

(vii) Goodness of Fit measures (R^2 & Adj R^2)

(ix) ANOVA Analysis.

(x) Testing of Hypothesis.

Violations of OLS Assumptions:-

Simple Model: $Y_i = \alpha + \beta X_i + u_i$.

Assumptions:

(i) $E(u_i) = 0 \quad \forall i$.

~~(ii)~~ $Var(u_i) = E(u_i^2) = \sigma^2 \; \forall i$

~~(iii)~~ $cov(u_i, u_j) = E(u_i u_j) = 0 \quad \forall \; i \neq j$

(iv) $cov(X, u) = 0$.

(v) $u_i \overset{iid}{\sim} N(0, \sigma^2)$ [Needs to incorporated only during Testing of Hypothesis].

If $Var(u_i) \neq \sigma^2 \; \forall i \Rightarrow$ Problem of Heteroskedasticity
If $cov(u_i, u_j) \neq 0 \; \forall \; i \neq j \Rightarrow$ Problem of Autocorrelation.

i.e from the statement of Gauss Markov Theorem, if Heteroskedasticity / Autocorrelation is present in the data, then the OLS estimates will no longer be BLUE.

Heteroskedasticity:

True Model: $Y_i = \alpha + \beta X_i + u_i$.

where: $\boxed{Var(u_i) = \sigma_i^2}$ [Rest Gauss Markov Assumptions hold].

Estimated Model: $\hat{Y}_i = \hat{\alpha}_{OLS} + \hat{\beta}_{OLS} X_i$

where $\hat{\beta}_{OLS} = \dfrac{\sum (Y_i - \bar{Y})(X_i - \bar{X})}{\sum (X_i - \bar{X})^2}$ $\quad \hat{\alpha}_{OLS} = \bar{Y} - \hat{\beta}_{OLS} \bar{X}$

$\therefore$ According to Gauss Markov Th, $\hat{\alpha}_{OLS}$ and $\hat{\beta}_{OLS}$ are not BLUE.

Suppose if we can express $\sigma_i^2 = \sigma^2 Z_i^2$.

$$\frac{Y_i}{Z_i} = \alpha\left(\frac{1}{Z_i}\right) + \beta \frac{X_i}{Z_i} + \frac{u_i}{Z_i}$$

Transformed Model: $\left(\dfrac{Y_i}{Z_i}\right) = \alpha\left(\dfrac{1}{Z_i}\right) + \beta\left(\dfrac{X_i}{Z_i}\right) + \left(\dfrac{u_i}{Z_i}\right)$ new random disturbance term.

$$\Rightarrow Var\left(\frac{u_i}{Z_i}\right) = \frac{1}{Z_i^2} Var(u_i) = \frac{\sigma_i^2}{Z_i^2} = \frac{\sigma^2 \cdot \cancel{Z_i^2}}{\cancel{Z_i^2}} = \sigma^2$$

$\therefore \left(\dfrac{u_i}{Z_i}\right)$ are Homoskedastic.

$\therefore$ As Transformed Model is Homoskedastic, applying OLS on the transformed model will give us efficient estimates of α, β.

This is the method of Weighted Least Squares (WLS).

WLS gives us estimates $\hat{\alpha}_{WLS}$ and $\hat{\beta}_{WLS}$.

According to Gauss-Markov Theorem, $\hat{\alpha}_{WLS}$ and $\hat{\beta}_{WLS}$ are efficient.

i.e, under Heteroskedasticity, $Var(\hat{\beta}_{WLS}) \leq Var(\hat{\beta}_{OLS})$

$$Var(\hat{\alpha}_{WLS}) \leq Var(\hat{\alpha}_{OLS})$$

Auto correlation:

More prominant is time-series data.

$cov(u_t, u_\tau) \neq 0$.

	X_t	Y_t
Pd 1		
Pd 2		
⋮		
Pd T		

True Model: $Y_t = \alpha + \beta X_t + u_t$.

$$\boxed{u_t = \rho . u_{t-1} + \varepsilon_t}, \quad \varepsilon_t \overset{iid}{\sim} N(0, \sigma_\varepsilon^2), \quad |\rho| < 1.$$

Dependence structure of u_t.

Auto correlation is present in the data and if OLS estimation is performed on the true model then $\hat{\alpha}_{OLS}$ and $\hat{\beta}_{OLS}$ will be ineffecient [From Gauss Markov Theorem].

True Model: Pd 't', $\quad Y_t = \alpha + \beta X_t + u_t$. ----- (i)

Pd 't-1' $\quad \rho Y_{t-1} = \alpha\rho + \beta.\rho X_{t-1} + \rho u_{t-1}$ --- (ii)

$$(i)-(ii): \left(Y_t - \rho . Y_{t-1}\right) = \alpha(1-\rho) + \beta\left(X_t - \rho . X_{t-1}\right) + \underbrace{\left(u_t - \rho . u_{t-1}\right)}_{= \varepsilon_t}$$

Transformed Model: $\quad (Y_t - \rho . Y_{t-1}) = \alpha(1-\rho) + \beta(X_t - \rho . X_{t-1}) + \varepsilon_t$.

Here $\varepsilon_t \sim$ Gauss Markov assumptions. Estimates of α, β. from this transformed Model will be effecient.

...tes of α, β.
from this transformed Model will be effecient.

Method of Generalized Least Squares (GLS).

GLS estimates: $\hat{\alpha}_{GLS}$, $\hat{\beta}_{GLS}$.

From Gauss Markov Theorem: $Var(\hat{\beta}_{GLS}) \leq Var(\hat{\beta}_{OLS})$

under Autocorrelation: $Var(\hat{\alpha}_{GLS}) \leq Var(\hat{\alpha}_{OLS})$