北京航空航天大學
经济管理学院
BEIHANG UNIVERSITY
SCHOOL OF ECONOMICS AND MANAGEMENT

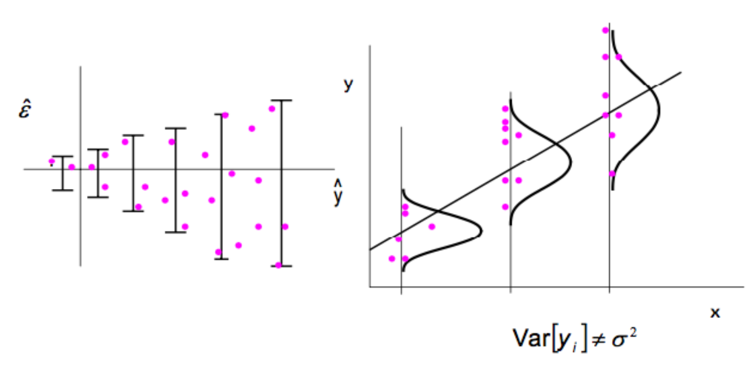
Generalized Linear Models

Lecture 6.0: Generalized Linear Squares

Outline

Introduction

- Think about the assumptions we made in linear regression.
- How about errors are not constant?



- Or not independent?
- *Generalized or weighted least squares* (GLS or WLS)

Outline

GLS

- Now we have the model

$$\boldsymbol{Y} = \boldsymbol{X}\beta + \epsilon \quad E(\epsilon) = 0 \quad \text{Var}(\epsilon) = \sigma^2\boldsymbol{V}$$

- GLS estimation can be obtained by matrix factorization $\boldsymbol{V} = \boldsymbol{K}\boldsymbol{K}$.
- We define

$$\boldsymbol{z} = \boldsymbol{K}^{-1}\boldsymbol{y} \quad \boldsymbol{B} = \boldsymbol{K}^{-1}\boldsymbol{X} \quad g = \boldsymbol{K}^{-1}\epsilon$$

We have

$$\boldsymbol{z} = \boldsymbol{B}\beta + \boldsymbol{g}, E(\boldsymbol{g}) = \boldsymbol{0}, \text{Var}(\boldsymbol{g}) = \sigma^2\boldsymbol{I}.$$

GLS

$$\mathbf{z} = \mathbf{B}\beta + \mathbf{g}$$

Now we can use OLS to estimate β:

$$\begin{aligned} S(\beta) &= (\mathbf{z} - \mathbf{B}\beta)'(\mathbf{z} - \mathbf{B}\beta) \\ &= (\mathbf{K}^{-1}\mathbf{y} - \mathbf{B}\beta)'(K^{-1}y - B\beta) \\ &= (\mathbf{y} - \mathbf{X}\beta)'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\beta) \end{aligned}$$

Take the derivative with respect to β and set it to $\mathbf{0}$, we get

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

The **GLS** estimator of β is

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

GLS in R

```r
nlme::gls(y ~ x, data = dat, weights = weights, ...)
```

Outline

WLS

- Some times the errors are uncorrelated, but have unequal variance. In this case we use *weighted least squares* (WLS).
- The covariance matrix of ϵ has the form:

$$\sigma^2 \mathbf{V} = \sigma^2 \begin{bmatrix} 1/w_1 & & & 0 \\ & 1/w_2 & & \\ & & \ddots & \\ 0 & & & 1/w_n \end{bmatrix}$$

- Let $\mathbf{W} = \mathbf{V}^{-1}$ with elements w_i, the **WLS** estimator of $\hat{\beta}$ is

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

WLS

- In WLS, observations with large variances get smaller weights than observations with smaller variances.
- Examples of possible weights are:
  - Error proportional to a predictor x_i suggests $w_i = x_i^{-1}$.
  - When an observation y_i is an average of several, n_i, observations at that point of the explanatory variable, then, $\text{Var}(y_i) = \sigma^2/n_i$ suggests $w_i = n_i$.

WLS in R

```r
lm(y ~ x, data = dat, weights = weights, ...)
```

Outline

IRWLS

- Sometimes we will have prior information on the weights w_i, others we might find, looking at residual plots, that the variability is a function of one or more explanatory variables.
- In these cases we have to estimate the weights, perform the analysis, re-estimate the weights again based on these results and perform the analysis again.
- This procedure is called *iteratively reweighted least squares* (IRWLS).

IRWLS example

Suppose $\text{Var}(\epsilon_i) = \gamma_0 + \gamma_1 x_1$:

1. Start with $w_i = 1$.
2. Use OLS to estimate β.
3. Use residuals to estimate γ, perhaps regress $\hat{\epsilon}^2$ on $\boldsymbol{x}_1$.
4. Recompute w_i and repeat go to step 2.