- 第三章中，我们假设了误差项同方差（MLR.5）：

$$\boxed{Var(u_i|x_{i1},x_{i2},\ldots,x_{ik})=\sigma^2} \qquad \sigma_i^2$$

- 基于MLR.5，我们计算OLS估计值的方差，建立了高斯-马尔科夫定理，提出了假设检验方法（t检验，F检验）
- 然而，同方差性并不总是成立（例子：储蓄与收入），这种情况称之为异方差性
- 本章将讨论出现异方差性时的一些修正措施

JEFFREY M. WOOLDRIDGE

Introductory Econometrics

A Modern Approach

SIXTH EDITION

Chapter 8 异方差性

章节框架

- 在这一章中，我们将探讨异方差性
- 首先，我们介绍异方差对OLS估计值性质的影响
- 之后，我们讨论异方差出现时的修正措施以及异方差性的检验方法
- 最后，我们提出异方差情况下最优的估计方法：加权最小二乘（GLS）

异方差性对OLS的影响

- **第三章中，考虑多元线性回归模型：**

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + u$$

- 通过最小化残差平方和，我们得到OLS估计值

$$\hat{\beta} = \left(\sum_{i=1}^{n} X_i X_i{}'\right)^{-1} \sum_{i=1}^{n} X_i y_i$$

- 异方差性(违反MLR.5)会影响OLS哪些性质？

异方差性对OLS的影响

- **OLS有限样本性质**
  - 期望的无偏性：MLR.1 – MLR.4 ✓
  - 方差公式：MLR.1 – MLR.5 同方差
  - 高斯-马尔科夫定理：MLR.1 – MLR.5

    方差最小。

    BLUE
  - 正态性：MLR.1 – MLR.6 正态分布。

σ_i^2

- **OLS大样本性质**
  - 一致性：MLR.1 – MLR.4′ ✓
  - 渐近正态性：MLR.1 – MLR.5
- **在异方差下，OLS仍然无偏且一致**
- **但是，异方差性可能导致其他性质不成立**
- **OLS不再是最优线性无偏估计量（BLUE）；可能会有其他更有效的线性估计量**

异方差性对OLS的影响

- **异方差性对OLS相关统计量的影响**

$\underline{\hat{\beta}} \leftarrow$? 数值没变.

$\hat{u}$, $\hat{y}$

- R方和调整R方的解释不受异方差性的影响：

$$R^2 \equiv SSE/SST = 1 - SSR/SST$$

$\hat{y} - \bar{y}$

- t统计量与F统计量在原假设下的分布依赖于OLS的方差形式与正态分布

- 因此，异方差性下t统计量与F统计量将失效

OLS估计后的异方差-稳健推断

- 已经研发出对于未知形式的异方差性稳健的OLS标准误差及其统计量
- <u>它们至少在大样本下是有效的</u>
- 异方差-稳健的OLS标准误差

Huber
Eiker
White (1980)

$$\widehat{Var}(\hat{\beta}_j) = \frac{\sum_{i=1}^{n} \hat{r}_{ij}^2 \hat{u}_i^2}{SSR_j^2}$$

也叫作<u>White/Huber/Eicker</u>标准误。其中包括原回归以及x_j对其他解释变量回归的残差平方和。

$e_j = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ 第j个

$\begin{bmatrix} 0 & & \\ & 0 & \\ & & 且j个 \\ & & & 0 \end{bmatrix}$

e_j' (夹心估计量) e_j

Huber- White

$$se(\hat{\beta}_j) = \sqrt{\widehat{Var}(\hat{\beta}_j)}$$

- 使用异方差-稳健的标准误重新构造t检验统计量：其中α_j 为假设值

"Robust" Standard error

Newey- West Standard error

$$t_{\hat{\beta}_j} = \frac{\hat{\beta}_j - \alpha_j}{se(\hat{\beta}_j)}$$

Martingale Difference Sequence M.D.S

- 异方差-稳健的t检验是渐近有效的
- 异方差-稳健的F统计量（瓦尔德统计量）

Wald

渐近独立 ergodicity

有限样本下的方差.

$$\underset{n\times k}{X}: \begin{bmatrix} x_{11} & x_{12} & & x_{1k} \\ \vdots & \vdots & \ddots & \\ x_{n1} & x_{n2} & & x_{nk} \end{bmatrix}$$

$$\hat{\beta} = \beta + \underbrace{(X'X)^{-1}X'}u$$

大样本时: $\underbrace{\sqrt{n}(\hat{\beta}-\beta)}$

$$\hat{u}_i^2$$

$$Var(\hat{\beta}|X) = (X'X)^{-1}X' \, Var(u|X) \, X(X'X)^{-1}$$

异方差:

$$= (X'X)^{-1}(X'\Omega X)(X'X)^{-1}, \quad \Omega = \begin{bmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_n^2 \end{bmatrix}$$

$$= (X'X)^{-1}\Big(\sum_{i=1}^{n} \sigma_i^2 x_i x_i'\Big)(X'X)^{-1}, \quad x_i = \begin{bmatrix} x_{i1} \\ \vdots \\ x_{iK} \end{bmatrix}$$

$K\times K$

K个自变量.

同方差时:

$$\sigma^2(X'X)^{-1}$$

估计值.

$$= (X'X)^{-1}\Big(\sum \hat{\sigma}_i^2 x_i x_i'\Big)(X'X)^{-1}$$

$$= (X'X)^{-1}\Big(\sum \hat{u}_i^2 x_i x_i'\Big)(X'X)^{-1}$$

核心, ①

$$\underbrace{e_j'(X'X)^{-1}\Big(\sum_{i=1}^{n} \hat{u}_i^2 x_i x_i'\Big)(X'X)^{-1} e_j}$$

第j个对角线上元素.

渐近方差

②式是①式的n倍

$$\hat{\beta} = \beta + (X'X)^{-1}X'u \qquad X'X = \sum_{i=1}^{n} x_i x_i'$$

$$\sqrt{n}\cdot(\hat{\beta}-\beta) = \sqrt{n}\left(\frac{1}{n}X'X\right)^{-1}\left(\frac{1}{n}X'u\right) = \left(\frac{1}{n}X'X\right)^{-1}\left(\frac{1}{\sqrt{n}}X'u\right)$$

CLT: $\frac{1}{\sqrt{n}}X'u \xrightarrow{d} N(0, Var(x_i u_i))$

$$E(u_i^2 x_i x_i') \approx \frac{1}{n}\sum \hat{\sigma}_i^2 x_i x_i'$$

$$Avar(\sqrt{n}(\hat{\beta}-\beta)) = E(x_i x_i')^{-1} E(u_i^2 x_i x_i') E(x_i x_i')^{-1}$$

同方差

$$\sigma^2 E(x_i x_i')^{-1}$$

$$\widehat{Avar}(\sqrt{n}(\hat{\beta}-\beta)) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i'\right)^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\hat{u}_i^2 x_i x_i'\right)\left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i'\right)^{-1}$$

$$= \left(\frac{1}{n}X'X\right)^{-1}\left(\frac{1}{n}X'\Omega X\right)\left(\frac{1}{n}X'X\right)^{-1} \quad ②$$

证明稳健协方差矩阵对角线元素可写成8.4式

$$\widehat{Var}(\hat{\beta}_j) = \sum_{i=1}^{n} \hat{u}_i^2 \left(\boxed{\underbrace{e_j'}_{1\times K}\underbrace{(X'X)^{-1}}_{K\times K}\underbrace{x_i}_{K\times 1}} \cdot \boxed{x_i'(X'X)^{-1}e_j} \right)$$

$$= \sum_{i=1}^{n} \hat{u}_i^2 \left(\underbrace{x_i'(X'X)^{-1}e_j} \right)^2$$

$$h_{ij} \equiv x_i'(X'X)^{-1}e_j \qquad \widehat{Var}(\hat{\beta}) = \sum_{i=1}^{n} \hat{u}_i^2 h_{ij}^2$$

$$(8.4)\ \widehat{Var}(\hat{\beta}) = \frac{\sum_{i=1}^{n} \hat{r}_{ij}^2 \hat{u}_i^2}{\left(\sum \hat{r}_{ij}^2\right)^2}$$，需证 $h_{ij} = \dfrac{\hat{r}_{ij}}{\sum \hat{r}_{ij}^2}$

挪动 x_j 到最后位置。

$$X'X = \begin{bmatrix} X_{-j}'X_{-j} & X_{-j}'x_j \\ x_j'X_{-j} & x_j'x_j \end{bmatrix}$$

$$X = [X_{-j} \quad x_j]$$

$$
(X'X)^{-1} = \begin{bmatrix} A^{-1} + A^{-1} b s^{-1} b' A^{-1} & -A^{-1} b s^{-1} \\ -s^{-1} b' A^{-1} & s^{-1} \end{bmatrix}
$$

Schur complement

$$
A = X'_{-j} X_{-j}, \quad b = X'_{-j} x_j, \quad c = x_j' x_j, \quad s = c - b' A^{-1} b .
$$

$$
s = x_j' x_j - x_j' \underline{X_{-j} (X'_{-j} X_{-j})^{-1} X'_{-j}} x_j = x_j' \underbrace{(I - P_{-j})}_{M_{-j}, \ M_{-j} \cdot M_{-j}} x_j = \hat{r}_j' \hat{r}_j
$$

$$
= \sum_{i=1}^{n} \hat{r}_{ij}^2
$$

$$
(X'X)^{-1} e_j, \ (\text{提最后一列}) = \begin{bmatrix} -A^{-1} b s^{-1} \\ s^{-1} \end{bmatrix}
$$

$$
x_i' = [\, x'_{i,-j} \quad x_{ij} \,] . \qquad x_i' (X'X)^{-1} e_j = -x'_{i,-j} A^{-1} b s^{-1} + x_{ij} s^{-1}
$$

$$
= s^{-1} \left(x_{ij} - x'_{i,-j} (X'_{-j} X_{-j})^{-1} X'_{-j} x_j \right)
$$

$$
\hat{r}_j = (I - P_{-j}) x_j = x_j - X_{-j} (X'_{-j} X_{-j})^{-1} X'_{-j} x_j
$$

故 $x_i'(X'X)^{-1}e_j = \dfrac{\hat{r}_{ij}}{S} = \dfrac{\hat{r}_{ij}}{\sum_{i=1}^{n} \hat{r}_{ij}^2}$

OLS估计后的异方差-稳健推断

- **例子：小时工资方程**

如果同方差，使用了Robust Std.
假设看到了显著性，则无所谓.

$$\widehat{\log}(wage) = -\ .128 + .0904\ educ + .0410\ exper - .0007\ exper^2$$
$$\begin{array}{cccc} & (.105) & (.0075) & (.0052) & (.0001) \\ & [.107] & [.0078] & [.0050] & [.0001] \end{array}$$

因为设为同方差

异方差稳健标准误与不稳健标准误看上去差不多

- （1） $H_0\!: \beta_{exper} = 0$

$t = 7.88 \quad t_{robust} = 8.2$

t统计量差不多

估计量，显著性
只会更强.
(BLUE)

- （2） $H_0 : \beta_{exper} = \beta_{exper^2} = 0$

$F = 17.95$

F统计量也差不多

$F_{robust} = 17.99$

如果异方差性强，差别就会大。

对异方差性的检验

- 我们总是关心OLS是否能保持BLUE，通常的t检验F检验是否有效
- 因此，我们想检验是否真的存在异方差性
- **同方差假设：**

$$\boxed{H_0 : Var(u|x_1, x_2, \ldots, x_k) = Var(u|\mathbf{x}) = \sigma^2}$$

$$\Rightarrow \quad E(u^2|x_1, \ldots, x_k) = E(u^2) = \sigma^2$$

- 同方差假设下u^2 的条件均值不随x_1，x_2，⋯，x_k 变化

对异方差性的检验

- 用残差的平方来对其他解释变量进行回归：

$$\hat{u}^2 = \delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k + error$$

- F检验：

$$H_0 : \delta_1 = \delta_2 = \cdots = \delta_k = 0$$

$$F = \frac{R^2_{\hat{u}^2}/k}{1 - R^2_{\hat{u}^2}/(n-k-1)}$$

- 布罗施-帕甘（Breusch-Pagan， BP）异方差检验：

$$LM = n \cdot R^2_{\hat{u}^2} \sim \chi^2_k$$

另一种检验统计量（拉格朗日乘数统计量，LM）。该统计量越高（高R2），越倾向于拒绝原假设

$F \sim F(q, n-k)$, $n \to \infty$, $qF \xrightarrow{d} \chi^2(q)$

对异方差性的检验

- **BP检验步骤：**

（1）计算OLS估计的残差平方$\hat{u}^2$

（2）进行回归$\hat{u}^2 = \delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k + error$ **并计算回归的R方:** $R^2_{\hat{u}^2}$

（3）利用$R^2_{\hat{u}^2}$**计算F 统计量或LM统计量的取值，并计算相应的p值，如果p值<显著性水平则拒绝同方差的原假设**

对异方差性的检验

- **例子：住房价格方程中的异方差性**

$$\widehat{price} = -\underset{(29.48)}{21.77} + \underset{(.0006)}{.0021}\ lotsize + \underset{(.013)}{.123}\ sqrft + \underset{(9.01)}{13.85}\ bdrms$$

异方差

$$\Rightarrow R^2_{\hat{u}^2} = .1601, \boxed{p\text{-}value_F = .002,\ p\text{-}value_{LM} = .0028}$$

$$\widehat{\log}(price) = -\underset{(.65)}{1.30} + \underset{(.038)}{.168}\ \log(lotsize) + \underset{(.093)}{.700}\ \log(sqrft) + \underset{(.028)}{.037}\ bdrms$$

$$\Rightarrow R^2_{\hat{u}^2} = .0480, \boxed{p\text{-}value_F = .245,\ p\text{-}value_{LM} = .2390}$$

在log形式下，同方差假设不能被拒绝

对异方差性的检验

- F检验和BP检验只考察了残差平方与解释变量的线性关系
- **White异方差检验**

用残差平方对解释变量、其平方及其交互项进行回归

$$\hat{u}^2 = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + \delta_3 x_3 + \delta_4 x_1^2 + \delta_5 x_2^2 + \delta_6 x_3^2$$

$$+ \delta_7 x_1 x_2 + \delta_8 x_1 x_3 + \delta_9 x_2 x_3 + error$$

$$H_0: \delta_1 = \delta_2 = \cdots = \delta_9 = 0$$

较BP检验来说，White检验考察更一般的异方差的偏差

$$LM = n \cdot R_{\hat{u}^2}^2 \sim \chi_9^2$$

k=3

6, 6, $\frac{6\times5}{2}$

- **White检验的缺点**
  - 估计参数过多（例如，若k=6，则有27个估计参数）

对异方差性的检验

- **White检验的特例：**

$$\hat{u}^2 = \delta_0 + \delta_1 \hat{y} + \delta_2 \hat{y}^2 + error$$

间接检验解释变量、其平方及其交互项对残差平方的影响，因为y的被预测值及其平方包含了以上解释变量及其变形。

$$H_0 : \delta_1 = \delta_2 = 0,\ LM = n \cdot R^2_{\hat{u}^2} \sim \chi^2_2$$

- **例子：对数住房价格方程中怀特检验的特殊形式**

$$R^2_{\hat{u}^2} = .0392, LM = 88(.0392) \approx 3.45, p\text{–}value_{LM} = .178$$

加权最小二乘

- 在特殊的异方差结构下，我们调整估计方法以获得最有效的估计
- 除了一个常数倍数以外异方差是已知的（ Heteroskedasticity is known up to a multiplicative constant ）

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + u_i$$

已知 h_i

$$\boxed{Var(u_i|\mathbf{x}_i) = \sigma^2 h(\mathbf{x}_i),\ h(\mathbf{x}_i) = h_i > 0}$$

除σ^2外异方差的函数形式是已知的

信息含量加权，第i组方差大，信息小。

- 变换模型，使得变换之后的模型满足同方差假设

$$\Rightarrow \left[\frac{y_i}{\sqrt{h_i}}\right] = \beta_0\left[\frac{1}{\sqrt{h_i}}\right] + \beta_1\left[\frac{x_{i1}}{\sqrt{h_i}}\right] + \cdots + \beta_k\left[\frac{x_{ik}}{\sqrt{h_i}}\right] + \left[\frac{u_i}{\sqrt{h_i}}\right]$$

$$\Leftrightarrow\ y_i^* = \beta_0 x_{i0}^* + \beta_1 x_{i1}^* + \cdots + \beta_k x_{ik}^* + u_i^*$$

加权最小二乘

- 变形后的模型是同方差的

$$E(u_i^{*2}|\mathbf{x}_i) = E\left[\left(\frac{u_i}{\sqrt{h_i}}\right)^2 |\mathbf{x}_i\right] = \frac{E(u_i^2|\mathbf{x})}{h_i} = \frac{\sigma^2 h_i}{h_i} = \sigma^2$$

- 例子：储蓄与收入

$$sav_i = \cancel{\beta_0} + \beta_1 inc_i + u_i, \;\; Var(u_i|inc_i) = \sigma^2 inc_i$$

$$\left[\frac{sav_i}{\sqrt{inc_i}}\right] = \beta_0\left[\frac{1}{\sqrt{inc_i}}\right] + \beta_1\left[\frac{inc_i}{\sqrt{inc_i}}\right] + u_i^*$$

注意：这个回
归没有截距项

- 如果其他的高斯-马尔科夫假定都成立，则对变形后的模型使用OLS是最优无偏线性估计

加权最小二乘

- **变形后模型的OLS回归等价于加权最小二乘**（weighted least squares,WLS）

$$\min \sum_{i=1}^{n}\left(\left[\frac{y_i}{\sqrt{h_i}}\right]-b_0\left[\frac{1}{\sqrt{h_i}}\right]-b_1\left[\frac{x_{i1}}{\sqrt{h_i}}\right]-\cdots-b_k\left[\frac{x_{ik}}{\sqrt{h_i}}\right]\right)^2$$

$$\Leftrightarrow \min \sum_{i=1}^{n}(y_i-b_0-b_1x_{i1}-\cdots-b_kx_{ik})^2/h_i$$

在优化问题中，对具有更大方差的观测值赋予更小的权重

- **WLS是广义最小二乘**(generalized least squares, GLS)**的一种特殊形式**

Generalized Method Moments,

GMM,

加权最小二乘

- **例子：金融财富方程**

金融财富净值

TABLE 8.1 Dependent Variable: *nettfa*

Independent Variables	(1) OLS	(2) WLS	(3) OLS	(4) WLS
inc	.821 (.104)	.787 (.063)	.771 (.100)	.740 (.064)
$(age - 25)^2$	—	—	.0251 (.0043)	.0175 (.0019)
male	—	—	2.48 (2.06)	1.84 (1.56)
e401k	—	—	6.89 (2.29)	5.19 (1.70)
intercept	−10.57 (2.53)	−9.58 (1.65)	−20.98 (3.50)	−16.70 (1.96)
Observations	2,017	2,017	2,017	2,017
R-squared	.0827	.0709	.1279	.1115

异方差的假设形式

WLS估计量具有明显更小的标准误，符合更有效的预期。

参与401K计划

加权最小二乘

- **例子：数据为平均值**
- **如果观测值是城市/县/州/国家/公司层面的均值，则应根据单位规模进行加权**

公司i的养老金计划平均投入量

公司i的平均收入和年龄

公司为投入匹配的数量

异方差的误差项

$$\overline{contrib}_i = \beta_0 + \beta_1 \overline{earns}_i + \beta_2 \overline{age}_i + \beta_3 mrate_i + \bar{u}_i$$

$$Var(\bar{u}_i|\boldsymbol{x}_i) = Var\left(\frac{1}{m_i}\sum_{e=1}^{m_i} u_{i,e}\,\middle|\boldsymbol{x}_i\right) = \frac{\sigma^2}{m_i}$$

如果在个人层面的误差是同方差时的误差方差

223页

如果误差在个人层面是同方差的，则应该使用等同于企业规模m_i权重的WLS。如果个人层面的同方差不是一定的，那可以在WLS后计算稳健的标准误（即对于变化后模型）。

加权最小二乘

- **模型化异方差形式（可行GLS，feasible GLS，FGLS）** h?

contrive

$$\boxed{Var(u|\mathbf{x}) = \sigma^2 \exp(\delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k) = \sigma^2 h(\mathbf{x})}$$

假设异方差的一般形式；指数函数用来保证为正

$$u^2 = \sigma^2 \exp(\delta_0 + \delta_1 x_1 + \cdots + \delta_k x_k) \cdot v$$

$E[v|x]=1$

倍数误差（假设：独立于解释变量）

$$\Rightarrow \ \log(u^2) = \alpha_0 + \delta_1 x_1 + \cdots + \delta_k x_k + e$$

$$\log(\hat{u}^2) = \hat{\alpha}_0 + \hat{\delta}_1 x_1 + \cdots + \hat{\delta}_k x_k + error$$

使用异方差函数估计值的倒数作为WLS的权重

$$\boxed{\Rightarrow \ \hat{h}_i = \exp(\hat{\alpha}_0 + \hat{\delta}_1 x_1 + \cdots + \hat{\delta}_k x_k)}$$

$\hat{y}, x_1^2, \sim \gamma_k x_k^2, x_1 x_2$

BP, White, Parker.

可行GLS是一致的，而且比OLS更加渐近有效

加权最小二乘

- **例子：对香烟的需求**
- **OLS估计**

在餐馆禁止抽烟
（虚拟变量）

每天抽烟数量

log的收入和香烟价格

$$\widehat{cigs} = -\underset{(24.08)}{3.64} + \underset{(.728)}{.880}\ \log(income) - \underset{(5.773)}{.751}\ \log(cigpric)$$

$$- \underset{(.167)}{.501}\ educ - \underset{(.160)}{.771}\ age - \underset{(.0017)}{.0090}\ age^2 - \underset{(1.11)}{2.83}\ restaurn$$

拒绝同方差性

$$n = 807, R^2 = .0526, p\text{-}value_{Breusch-Pagan} = .000$$

加权最小二乘

- **FGLS估计**

现在统计上显著了

$$\widehat{cigs} = -\underset{(17.80)}{5.64} + \boxed{\underset{(.44)}{1.30}}\log(income) - \underset{(4.46)}{2.94}\log(cigpric)$$

$$-\underset{(.120)}{.463}\,educ + \underset{(.097)}{.482}\,age - \underset{(.0009)}{.0056}\,age^2 - \underset{(.80)}{3.46}\,restaurn$$

$$n = 807, R^2 = .1134$$

- **讨论**
  - 收入弹性现在是统计显著的

加权最小二乘

实践中

- <u>如果异方差函数错了会怎样？</u>

无解　　$\hat{u}^2 \to \hat{\sigma}^2$

① Robust, 显著

  - 如果异方差函数被错误设定，在MLR.1 – MLR.4下，WLS仍然是一致的，但是应该计算稳健标准误
  - WLS在MLR.4下是一致的，但在MLR.4’下却不一定一致

$\hat{u} \to u$?

$E[u|x]=0$

② Robust 不显著.

$$E(u_i|\mathbf{x}_i) = 0 \quad \Rightarrow \quad E\left(\left(u_i/\sqrt{h(\mathbf{x}_i)}\right)|\mathbf{x}_i\right) = 0$$

$cov(x,u)=0$?

选项很多

  - 如果OLS和WLS的估计值差别很大，这意味着可能一些假定（例如MLR.4）错了

充足理由

  - 如果异方差函数被错误设定，那理论上不能保证WLS比OLS更有效
  - （注意：经验上讲，如果具有很强的异方差性，使用一个错误的异方差形式来提高有效性也是不错的选择）

再议线性概率模型

- **在线性概率模型（LPM）中的WLS**

$$P(y=1|\mathrm{x}) = p(\mathrm{x}) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k$$

$$\Rightarrow \quad Var(y|\mathrm{x}) = p(\mathrm{x})\left[1 - p(\mathrm{x})\right]$$

在LPM中，异方差的准确形式是知道的

$$\Rightarrow \quad \widehat{h}_i = \widehat{y}_i(1 - \widehat{y}_i)$$

在WLS中使用倒数来作为权重

- **讨论**
  - 如果LPM的预测值小于0或大于1，则WLS不可行
  - 如果这种情况很少见，可以将其调整为0.01/0.99这种值
  - 否则，最好使用考虑稳健标准误的OLS回归

contrive.

Robust. Error.

默认选项

本章小节

- 异方差性虽然不会导致 OLS 估计量的偏误和不一致性，却致通常的标准误和检验统计量都不再成立
- 我们说明了如何计算异方差一稳健的标准误和 t 统计量
- 我们讨论了检验异方差性的两种常见方法:布罗施-帕甘检验和怀特检验
- 出现异方差性时，OLS 就不再是最优线性无偏的估计量。若知道异方差性的形式，则可使用WLS/GLS
- 更常见的是，我们在应用 WLS之前必须先估计一个异方差模型，由此得到可行的 GLS （FGLS）估计量

但还可能存在比 OLS 更有效的方法，比如 GLS。

2. 广义最小二乘法(GLS)

假设$\mathrm{Var}(\boldsymbol{\varepsilon}\mid \boldsymbol{X})=\sigma^2\boldsymbol{V}(\boldsymbol{X})\neq\sigma^2\boldsymbol{I}_n$，其中$\boldsymbol{V}(\boldsymbol{X})$为对称正定矩阵且已知，可能依赖于$\boldsymbol{X}$。

GLS 的基本思想是，通过变量转换，使得转换后的模型满足球型扰动项的假定。

命题 对于对称正定矩阵$\boldsymbol{V}_{n\times n}$，存在非退化矩阵$\boldsymbol{C}_{n\times n}$，使得$\boldsymbol{V}^{-1}=\boldsymbol{C}'\boldsymbol{C}$。

在一维情况下，“$\boldsymbol{V}$ 正定”即要求$\boldsymbol{V}$为正数，故$\dfrac{1}{\boldsymbol{V}}$也是正数，可分解为$\dfrac{1}{\sqrt{\boldsymbol{V}}}\cdot\dfrac{1}{\sqrt{\boldsymbol{V}}}$；如果$\boldsymbol{V}$为负数，则无法进行此分解。

矩阵$\boldsymbol{C}$不唯一，但不影响 GLS 的最终结果。

将原回归模型$\boldsymbol{y}=\boldsymbol{X\beta}+\boldsymbol{\varepsilon}$两边同时左乘矩阵$\boldsymbol{C}$：

$$\boldsymbol{Cy}=\boldsymbol{CX\beta}+\boldsymbol{C\varepsilon}$$

定义变量转换：

$$\tilde{\boldsymbol{y}}\equiv\boldsymbol{Cy},\ \tilde{\boldsymbol{X}}\equiv\boldsymbol{CX},\ \tilde{\boldsymbol{\varepsilon}}\equiv\boldsymbol{C\varepsilon}$$

可将模型写为

$$\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{X}}\boldsymbol{\beta} + \tilde{\boldsymbol{\varepsilon}}$$

变换后的模型仍满足严格外生性：

$$\mathrm{E}(\tilde{\boldsymbol{\varepsilon}} \mid \tilde{\boldsymbol{X}}) = \mathrm{E}(\boldsymbol{C}\boldsymbol{\varepsilon} \mid \boldsymbol{C}\boldsymbol{X}) = \mathrm{E}(\boldsymbol{C}\boldsymbol{\varepsilon} \mid \boldsymbol{X}) = \boldsymbol{C}\,\mathrm{E}(\boldsymbol{\varepsilon} \mid \boldsymbol{X}) = \boldsymbol{0}$$

球型扰动项的假定也得到满足：

$$\begin{aligned}\mathrm{Var}(\tilde{\boldsymbol{\varepsilon}} \mid \tilde{\boldsymbol{X}}) &= \mathrm{E}(\tilde{\boldsymbol{\varepsilon}}\,\tilde{\boldsymbol{\varepsilon}}' \mid \boldsymbol{X}) = \mathrm{E}(\boldsymbol{C}\boldsymbol{\varepsilon}\,\boldsymbol{\varepsilon}'\boldsymbol{C}' \mid \boldsymbol{X}) = \boldsymbol{C}\,\mathrm{E}(\boldsymbol{\varepsilon}\,\boldsymbol{\varepsilon}' \mid \boldsymbol{X})\boldsymbol{C}' = \sigma^2\boldsymbol{C}\boldsymbol{V}\boldsymbol{C}' \\ &= \sigma^2\boldsymbol{C}(\boldsymbol{V}^{-1})^{-1}\boldsymbol{C}' = \sigma^2\boldsymbol{C}(\boldsymbol{C}'\boldsymbol{C})^{-1}\boldsymbol{C}' = \sigma^2\boldsymbol{C}\boldsymbol{C}^{-1}(\boldsymbol{C}')^{-1}\boldsymbol{C}' = \sigma^2\boldsymbol{I}_n\end{aligned}$$

故高斯–马尔可夫定理成立。

对变换后的模型使用 OLS 即得到 GLS 估计量：

$$\hat{\boldsymbol{\beta}}_{\mathrm{GLS}} = (\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{X}})^{-1}\tilde{\boldsymbol{X}}'\tilde{\boldsymbol{y}} = \left[(\boldsymbol{CX})'(\boldsymbol{CX})\right]^{-1}(\boldsymbol{CX})'\boldsymbol{Cy}$$

$$= (\boldsymbol{X}'\boldsymbol{C}'\boldsymbol{CX})^{-1}\boldsymbol{X}'\boldsymbol{C}'\boldsymbol{Cy} = (\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$$

虽然$\boldsymbol{C}$不唯一，但$\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}$唯一，因为$\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}$不依赖于$\boldsymbol{C}$。

由于高斯–马尔可夫定理成立，故$\hat{\boldsymbol{\beta}}_{\mathrm{GLS}}$是 BLUE，比 OLS 更有效。但前提是必须知道协方差矩阵$\boldsymbol{V}$。