



金融计量学

(Financial Econometrics)

主讲教师：任 飞

商学院金融系

办公地点：商学院1601

email: olrenfei@163.com



第三章 异方差和自相关

第一节 异方差的介绍

- 一、异方差的定义及产生原因
- 二、异方差的后果

第二节 异方差的检验

- 一、图示法
- 二、解析法

第三节 异方差的修正

- 一、当 $\sigma_{\varepsilon_t}^2$ 为已知
- 二、当 $\sigma_{\varepsilon_t}^2$ 为未知
- 三、模型对数变换法

第四节 金融实例分析

第五节 自相关的概念和产生原因

- 一、滞后值与自相关的概念
- 二、自相关产生的原因

第六节 自相关的度量和后果

- 一、自相关的度量
- 二、出现自相关的后果

第七节 自相关的检验和修正

- 一、自相关的检验方法
- 二、自相关的修正方法



第三章 异方差和自相关

第一节 异方差的介绍

一、异方差的定义及产生原因

二、异方差的后果



一、异方差的定义及产生原因

定义：

异方差（heteroscedasticity）：如果线性模型中随机误差项 ε_i 的方差 $var(\varepsilon_i)$ 的数值对不同的样本观察值各不相同，则称 ε_i 具有异方差，即：

$$var(\varepsilon_i) = E(\varepsilon_i^2) = \sigma_{\varepsilon_i}^2 \neq \text{常数}, i = 1, 2, \dots, n \quad (3.2)$$

产生原因：

- 随机误差项 ε 包括了测量误差和模型中被省略的一些因素对因变量的影响。
- 不同抽样单元的因变量观察值之间可能差别很大。因此异方差性多出现在横截面样本之中。（例如：浦发银行和贵州茅台的股票价格）



二、异方差的后果

(一) 线性和无偏性

α 、 β 的OLS估计量通过整理可得：

$$\begin{aligned}\hat{\alpha} &= \alpha + \sum \left(\frac{1}{n} - \bar{X}k_i \right) \varepsilon_i \\ \hat{\beta} &= \beta + \sum k_i \varepsilon_i\end{aligned}\tag{3.3}$$

其中， $k_i = \frac{x_i}{(\sum x_i^2)}$ ， $x_i = X_i - \bar{X}$

显然 $\hat{\alpha}$ 和 $\hat{\beta}$ 是 ε_i 的线性函数，**线性成立**。



二、异方差的后果

(一) 线性和无偏性

同时有：

$$E(\hat{\alpha}) = \alpha + \sum \left(\frac{1}{n} - \bar{x}k_i \right) E(\varepsilon_i) = \alpha \quad (3.4)$$

$$E(\hat{\beta}) = \beta + \sum k_i E(\varepsilon_i) = \beta$$

上式的成立与 ε_i 是否具有异方差无关，**无偏性成立**。



二、异方差的后果

(二) 最优性

$$\text{var}(\hat{\beta}) = \frac{\sum x_i^2 \sigma_i^2}{(\sum x_i^2)^2} \quad (3.5)$$

- 若同方差假定成立, $\text{var}(\varepsilon_i) = \sigma_i^2 = \sigma^2$, $\text{var}(\hat{\beta}) = \frac{\sigma^2 \sum x_i^2}{(\sum x_i^2)^2} = \frac{\sigma^2}{\sum x_i^2}$
- 若 ε_i 存在异方差, 假设方差随 X_i^2 变化, $\text{var}(\varepsilon_i) = \sigma_i^2 = \sigma^2 X_i^2$,

记异方差情况下 β 的OLS估计为 $\bar{\beta}$,

$$\text{var}(\bar{\beta}) = \frac{\sigma^2 \sum x_i^2 X_i^2}{(\sum x_i^2)^2} = \frac{\sigma^2}{\sum x_i^2} \cdot \frac{\sum x_i^2 X_i^2}{\sum x_i^2} = \text{var}(\hat{\beta}) \cdot \frac{\sum x_i^2 X_i^2}{\sum x_i^2} \quad (3.6)$$

多数情况下, $\frac{\sum x_i^2 X_i^2}{\sum x_i^2} > 1$, $\text{var}(\bar{\beta}) > \text{var}(\hat{\beta})$, **最优性不成立。**

$\bar{\beta}$ 的方差非最小, 使得 y 的预测值精度降低, OLS法已不再适用。



第三章 异方差和自相关

第二节 异方差的检验

一、图示法

二、解析法

一、图示法

(一) 因变量 y 与解释变量 x 的散点图

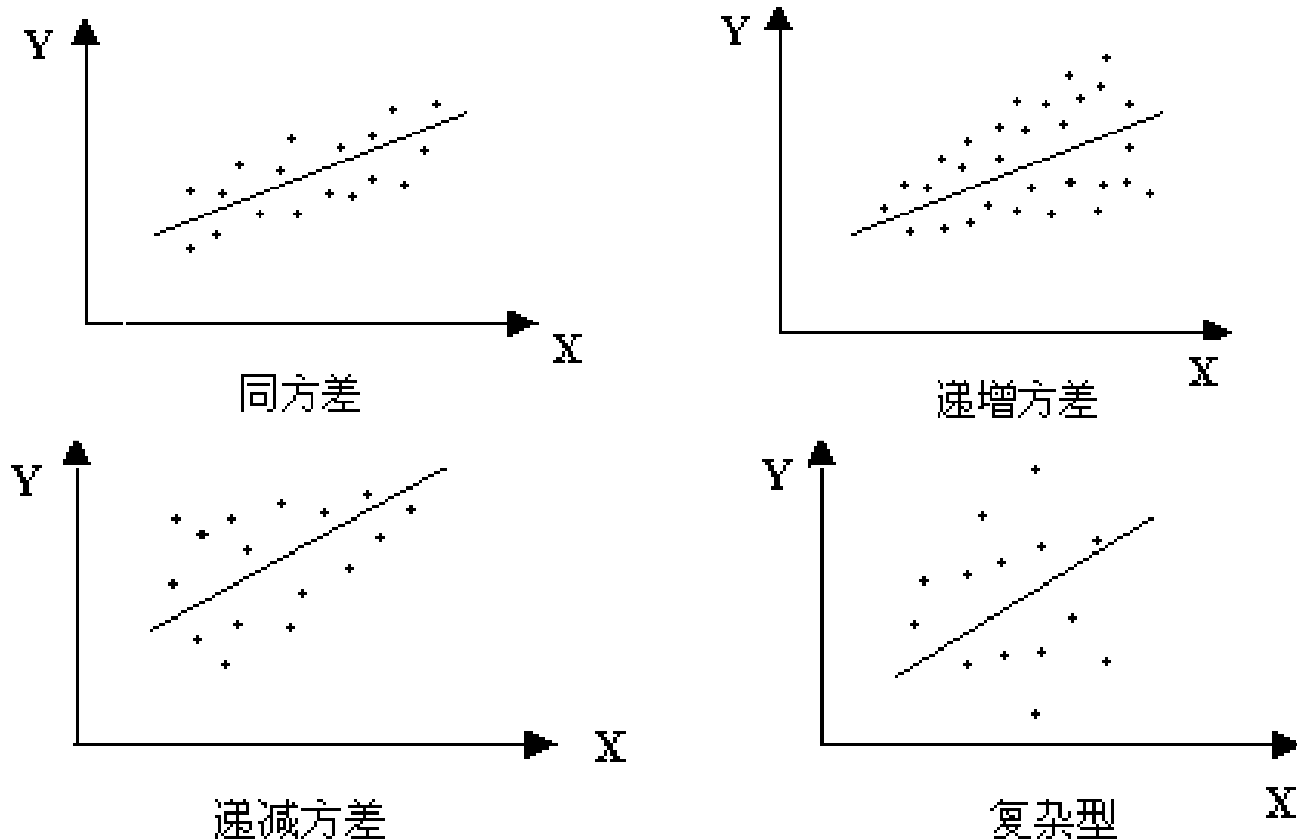


图3-1 因变量 y 与解释变量 x 的散点图

一、图示法

(二) 残差图

残差图即残差平方 $\hat{\varepsilon}_i^2$ 与 x_i 或 \hat{y}_i 的散点图。

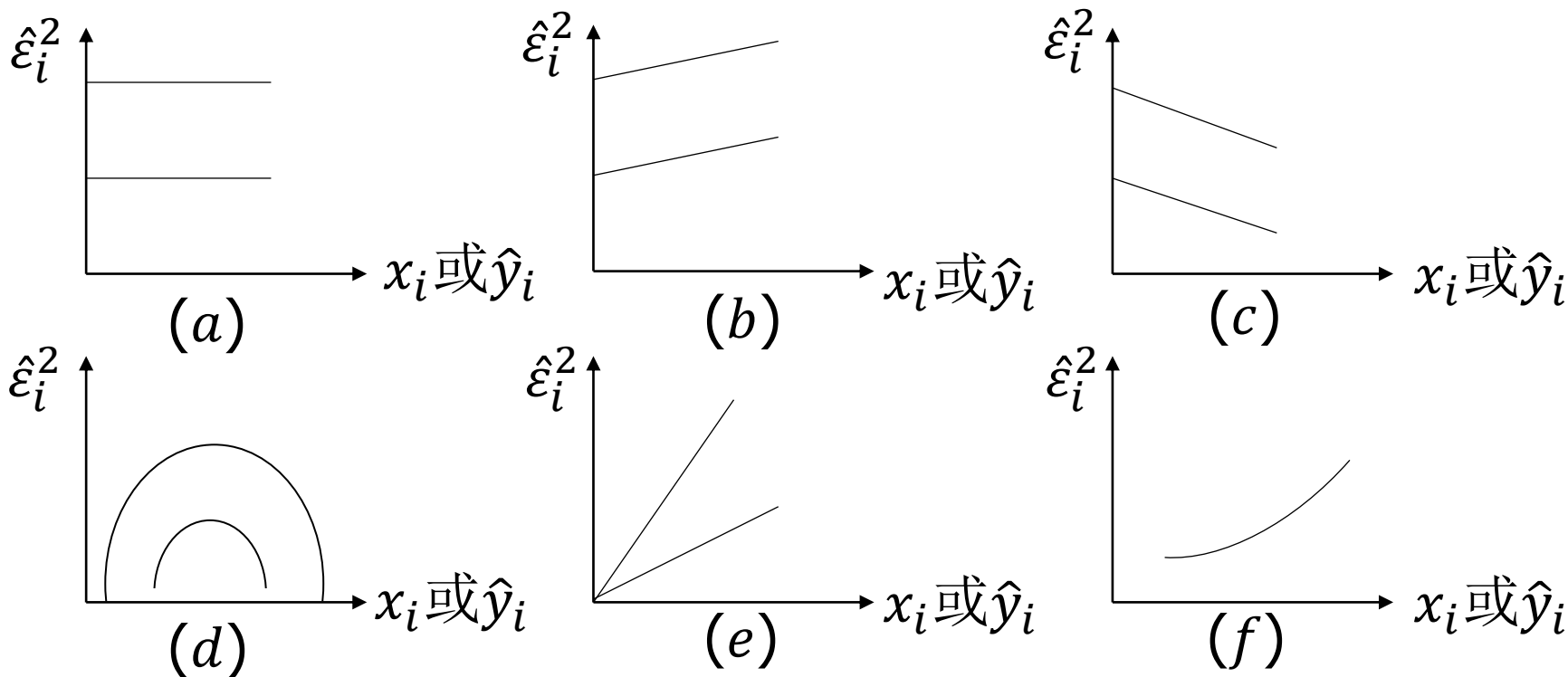


图3-2 回归模型的残差图



二、解析法

检验异方差的解析方法的共同思想是，由于不同的观察值随机误差项具有不同的方差，因此检验异方差的主要问题是**判断随机误差项的方差与解释变量之间的相关性**，下列这些方法都是围绕这个思路，通过建立不同的模型和验判标准来检验异方差：

- （一）Goldfeld-Quandt检验法
- （二）Spearman rank correlation检验法
- （三）Park检验法
- （四）White检验



二、解析法

(一) Goldfeld-Quandt检验法

- 以F检验为基础，适用于大样本情形（ $n > 30$ ）。
- 要求满足条件：（1）观测值的数目至少是参数的二倍；
（2）随机项没有自相关并且服从正态分布。
- 统计假设： $H_0: \varepsilon_i$ 是同方差（ $i = 1, 2, \dots, n$ ）
 $H_1: \varepsilon_i$ 具有异方差



二、解析法

(一) Goldfeld-Quandt检验法

➤ 具体步骤

第一步，处理观测值。

将某个解释变量的观测值按由小到大的顺序排列，然后将居中的 d 项观测数据除去，其中 d 的大小可以选择，比如取样本容量的 $1/4$ 。再将剩余的 $(n - d)$ 个数据分为数目相等的二组。

第二步，建立回归方程求残差平方和。

拟合两个回归模型，第一个是关于较小 x 值的那部分数据，第二个是关于较大 x 值的那部分数据。

对每一个回归模型，计算残差平方和：记 x_i 值较小的一组子样本的残差平方和为 $RSS_1 = \sum \varepsilon_{1i}^2$ ， x_i 值较大的一组子样本的残差平方和为 $RSS_2 = \sum \varepsilon_{2i}^2$ 。



二、解析法

(一) Goldfeld-Quandt检验法

➤ 具体步骤

第三步，建立统计量。

用所得出的两个子样本的残差平方和构成 F 统计量，若 H_0 为真，则：

$$F = \frac{\sum \varepsilon_{2i}^2 / (\frac{n-d}{2} - k - 1)}{\sum \varepsilon_{1i}^2 / (\frac{n-d}{2} - k - 1)} = \frac{\sum \varepsilon_{2i}^2}{\sum \varepsilon_{1i}^2} \sim F(\frac{n-d}{2} - k - 1, \frac{n-d}{2} - k - 1)$$

其中， n 为样本容量（观测值总数）， d 为被去掉的观测值数目， k 为模型中自变量的个数。



二、解析法

（一）Goldfeld-Quandt检验法

➤ 具体步骤

第四步，得出结论。

假设随机项服从正态分布（并且不存在序列相关），则统计量 RSS_2/RSS_1 将服从分子自由度和分母自由度均为 $(\frac{n-d}{2} - k - 1)$ 的 F 分布。

对于给定的显著性水平，如果统计量的值大于上述 F 分布的临界值，就拒绝原假设，认为残差具有异方差性。否则，不拒绝原假设。



二、解析法

(二) Spearman rank correlation检验法

假设模型为：

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

➤ 具体步骤

第一步，运用OLS法对原方程进行回归，计算残差 $\hat{\varepsilon}_i = y_i - \hat{y}_i$ ， $i = 1, 2 \dots n$ 。



二、解析法

(二) Spearman rank correlation检验法

第二步，计算Spearman等级相关系数。

将 $|\hat{\varepsilon}_i|$ 和解释变量观察值 x_i 按从小到大或从大到小的顺序分成等级。等级的大小可以人为规定，一般取大小顺序中的序号。如有两个值相等，则规定这个值的等级取相继等级的算术平均值。

然后，计算 $|\hat{\varepsilon}_i|$ 与 x_i 的等级差 d_i ， $d_i = x_i$ 的等级 - $|\hat{\varepsilon}_i|$ 的等级。最后根据公式计算Spearman等级相关系数。

$$r_s = 1 - 6 \left[\frac{\sum d_i^2}{n(n^2 - 1)} \right]$$

其中 d_i 表示第 i 个单元或现象的两种不同特性所处的等级之差，而 n 表示带有级别的单元或现象的个数。



二、解析法

(二) Spearman rank correlation 检验法

第三步，对总体等级相关系数 ρ_s 进行显著性检验 $H_0: \rho_s = 0$,
 $H_1: \rho_s \neq 0$ 。

样本 r_s 的显著性可通过t检验按下述方法加以检验：

$$t = \frac{r_s \sqrt{n-2}}{\sqrt{1-r_s^2}} \sim t(n-2)$$

对给定的显著水平 α ，查t分布表得 $t_{\frac{\alpha}{2}}(n-2)$ 的值，若 $|t| > t_{\frac{\alpha}{2}}(n-2)$ ，
表明样本数据异方差性显著，否则，认为不存在异方差性。



二、解析法

(三) Park检验法

- Park检验法就是将残差图法公式化，提出 $\sigma_{\varepsilon_i}^2$ 是解释变量 x_i 的某个函数，然后通过检验这个函数形式是否显著，来判定是否具有异方差性及其异方差性的函数结构。



二、解析法

(三) Park检验法

➤ 主要步骤

第一步，建立被解释变量 y 对所有解释变量 x 的回归方程，然后计算残差 $\hat{\varepsilon}_i^2$ ($i = 1, 2, \dots, n$)。

第二步，取异方差结构的函数形式为 $\sigma_{\varepsilon_i}^2 = \sigma^2 x_i^\beta e^{v_i}$ ，其中 σ^2 和 β 是两个未知参数， v_i 是随机变量。写成对数形式则为：

$$\ln \sigma_{\varepsilon_i}^2 = \ln \sigma^2 + \beta \ln x_i + v_i$$

第三步，建立方差结构回归模型，同时用 $\hat{\varepsilon}_i^2$ 来代替 $\sigma_{\varepsilon_i}^2$ ，即 $\ln \hat{\varepsilon}_i^2 = \ln \sigma^2 + \beta \ln x_i + v_i$ 。对此模型运用OLS法。对 β 进行t检验，如果不显著，则没有异方差性。否则表明存在异方差。



二、解析法

（四）White检验

➤ White检验不需要关于异方差的任何先验知识，只要求在大样本的情况下。

➤ 主要步骤

设二元线性回归模型为

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \varepsilon_t \quad (3.12)$$

异方差与解释变量的一般线性关系为

$$\sigma_{\varepsilon_t}^2 = \alpha_0 + \alpha_1 X_{2t} + \alpha_2 X_{3t} + \alpha_3 X_{2t}^2 + \alpha_4 X_{3t}^2 + \alpha_5 X_{2t} X_{3t} + v_t \quad (3.13)$$



二、解析法

(四) White检验

➤ 主要步骤

第一步，用OLS法估计式(3.12)的参数。

第二步，计算残差序列 $\hat{\varepsilon}_t$ 和 $\hat{\varepsilon}_t^2$ 。

第三步，求 $\hat{\varepsilon}_t^2$ 对 X_{2t} , X_{3t} , X_{2t}^2 , X_{3t}^2 , $X_{2t}X_{3t}$ 的线性回归估计式，即构造辅助回归函数。

第四步，计算统计量 nR^2 ，其中 n 为样本容量， R^2 为辅助回归函数中的决定系数。

第五步，在的 $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_5 = 0$ 原假设下， nR^2 服从自由度为5的 χ^2 分布，给定显著性水平 α ，查分布表得临界值 $\chi_\alpha^2(5)$ ，比较 nR^2 与 $\chi_\alpha^2(5)$ ，如果 $nR^2 > \chi_\alpha^2(5)$ ，则拒绝原假设，表明式(3.12)中随机误差存在异方差。



第三章 异方差和自相关

第三节 异方差的修正

一、当 $\sigma_{\varepsilon_i}^2$ 为已知

二、当 $\sigma_{\varepsilon_i}^2$ 为未知

三、模型对数变换法



第三节 异方差的修正

- 异方差性虽然不损坏OLS估计量的无偏性和一致性，但却使它们**不再是有用的**，甚至不是渐近（即在大样本中）有效的。**参数的显著性检验失效，降低了预测精度**。故而直接运用普通最小二乘法进行估计不再是恰当的，需要采取相应的修正补救办法以克服异方差的不利影响。
- 基本思路：变异方差为同方差，或者尽量缓解方差变异的程度。
- 将会遇到的情形分为两种：当误差项方差为 $\sigma_{\varepsilon_i}^2$ 已知和当 $\sigma_{\varepsilon_i}^2$ 未知。



一、异方差的修正-当 $\sigma_{\varepsilon_i}^2$ 为已知

加权最小二乘法（weighted least squares, WLS）

➤ **基本思想：**对较大的残差平方赋予较小的权数，对较小的残差平方赋予较大的权数。对残差所提供信息的重要程度作一番校正，以提高参数估计的精度。

可以考虑用 $\frac{1}{\text{var}(u_i)} = \frac{1}{\sigma_{\varepsilon_i}^2}$ 作为 ε_i^2 的权数，使加权残差平方和 $\sum \frac{\hat{\varepsilon}_i^2}{\sigma_{\varepsilon_i}^2} =$

$\sum \frac{1}{\sigma_{\varepsilon_i}^2} (y_i - \hat{\alpha} - \hat{\beta}x_i)^2$ 最小。

例：设异方差是 x_i 的函数： $\sigma_{\varepsilon_i}^2 = k^2 f(x_i)$ ，

加权最小二乘法即要求 $\sum \frac{\hat{\varepsilon}_i^2}{\sigma_{\varepsilon_i}^2} = \sum \frac{(y_i - \hat{\alpha} - \hat{\beta}x_i)^2}{k^2 f(x_i)}$ 最小

已知方差的情形在计量经济研究中很少发生。



二、异方差的修正-当 $\sigma_{\varepsilon_i}^2$ 为未知

现实中多数情况下 $\sigma_{\varepsilon_i}^2$ 是未知的，所以还要考虑别的方法来消除异方差。一般来讲，可以将异方差的表现分为：

- （一） $\sigma_{\varepsilon_i}^2$ 正比于 X_i^2 ；
- （二） $\sigma_{\varepsilon_i}^2$ 正比于 X_i ；
- （三） $\sigma_{\varepsilon_i}^2$ 正比于 Y 均值的平方

以 $Y_i = \alpha + \beta X_i + \varepsilon_i$ 为模型演示。



二、异方差的修正-当 $\sigma_{\varepsilon_i}^2$ 为未知

(一) $\sigma_{\varepsilon_i}^2$ 正比于 X_i^2

$$E(\varepsilon_i^2) = \sigma^2 X_i^2$$

可对原方程做如下变换：

$$\frac{Y_i}{X_i} = \frac{\alpha}{X_i} + \beta + \frac{\varepsilon_i}{X_i} = \alpha \frac{1}{X_i} + \beta + v_i$$

$E(v_i^2) = E\left(\frac{\varepsilon_i}{X_i}\right)^2 = \frac{1}{X_i^2} E(\varepsilon_i^2) = \sigma^2$ ， v_i 的方差是同方差性的，故可对变换方程施行OLS。



二、异方差的修正-当 $\sigma_{\varepsilon_i}^2$ 为未知

(二) $\sigma_{\varepsilon_i}^2$ 正比于 X_i

$$E(\varepsilon_i^2) = \sigma^2 X_i$$

可对原方程做如下变换：

$$\frac{Y_i}{\sqrt{X_i}} = \frac{\alpha}{\sqrt{X_i}} + \beta\sqrt{X_i} + \frac{\varepsilon_i}{\sqrt{X_i}} = \alpha \frac{1}{\sqrt{X_i}} + \beta\sqrt{X_i} + v_i$$

$$E(v_i^2) = E\left(\frac{\varepsilon_i}{\sqrt{X_i}}\right)^2 = \frac{1}{X_i} E(\varepsilon_i^2) = \sigma^2, \quad v_i \text{ 的方差是同方差性的。}$$



二、异方差的修正-当 $\sigma_{\varepsilon_i}^2$ 为未知

(三) $\sigma_{\varepsilon_i}^2$ 正比于 Y 均值的平方

$$E(\varepsilon_i^2) = \sigma^2 [E(Y_i)]^2$$

可对原方程做如下变换:

$$\frac{Y_i}{E(Y_i)} = \frac{\alpha}{E(Y_i)} + \beta \frac{X_i}{E(Y_i)} + \frac{\varepsilon_i}{E(Y_i)} = \alpha \frac{1}{E(Y_i)} + \beta \frac{X_i}{E(Y_i)} + v_i$$

$$E(v_i^2) = E\left(\frac{\varepsilon_i}{E(Y_i)}\right)^2 = \frac{1}{[E(Y_i)]^2} E(\varepsilon_i^2) = \sigma^2, \quad v_i \text{ 的方差是同方差性的。}$$

由于 Y_i 依赖 α 和 β , 一般先对 $Y_i = \alpha + \beta X_i + \varepsilon_i$ 进行**OLS**回归, 估计得到 $\hat{Y}_i = \hat{\alpha} + \hat{\beta} X_i$, 作为对 $E(Y_i)$ 的估计, 再做如下变换:

$$\frac{Y_i}{\hat{Y}_i} = \frac{\alpha}{\hat{Y}_i} + \beta \frac{X_i}{\hat{Y}_i} + \frac{\varepsilon_i}{\hat{Y}_i}$$

当样本量合理地大时, 上述变换会得到比较好的结果。



二、异方差的修正-当 $\sigma_{\varepsilon_i}^2$ 为未知

在上述变换中，都可以看到对 $\sigma_{\varepsilon_i}^2$ 的形式采取的是一种猜测的态度，即我们也不能肯定采取哪种变换更有效。

同时这些变换可能还有其他的一些问题：

（1）当解释变量多于1个时，也许先验上不知道应选择哪一个 X 去进行变换；

（2）当 $\sigma_{\varepsilon_i}^2$ 无法直接得知而要做出估计时，所有用到t检验、F检验等的检验程序，都只有在**大样本**中有效。因此，**在小的或有限的样本**中，应谨慎根据各种变换解释所得到的结果。



三、模型对数变换法

对模型 $Y_i = \alpha + \beta X_i + \varepsilon_i$ 中的变量取对数，得到：

$$\ln Y_i = \alpha + \beta \ln X_i + \varepsilon_i$$

对上述模型进行估计，通常可以降低异方差性的影响。

原因：

- （1）对数变换能够使测定变量值的尺度缩小；
- （2）经过对数变换的线性模型其残差 ε_i 表示相对误差，而相对误差往往具有较小的差异。



第三章 异方差和自相关

第四节 金融实例分析



第四节 金融实例分析

下面通过一个具体金融案例来讨论异方差的检验与修正过程：

例3.1 根据**1985-2019**年上海市城镇居民人均储蓄与人均可支配收入的数据资料，若假定 $income$ 为人均收入（元）， $saving$ 为人均储蓄（元），分析人均储蓄受人均收入的线性影响，可建立一元线性回归模型进行分析。

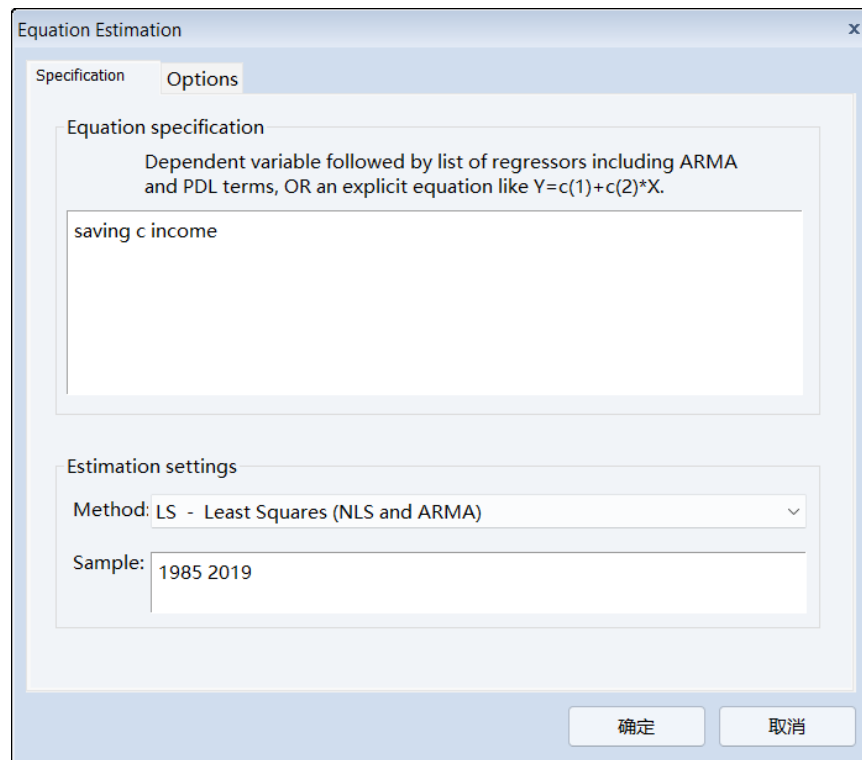
设模型为：

$$saving = \beta_1 + \beta_2 income + \mu$$



第四节 金融实例分析

1. 用OLS估计法估计参数：“Quick” → “Estimate Equation”



Equation Estimation

Specification Options

Equation specification

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)*X$.

saving c income

Estimation settings

Method: LS - Least Squares (NLS and ARMA)

Sample: 1985 2019

确定 取消



第四节 金融实例分析

1. 用OLS估计法估计参数

表3-1 Eviews回归结果

Dependent Variable: SAVING

Method: Least Squares

Date: 03/22/22 Time: 13:33

Sample: 1985 2019

Included observations: 35

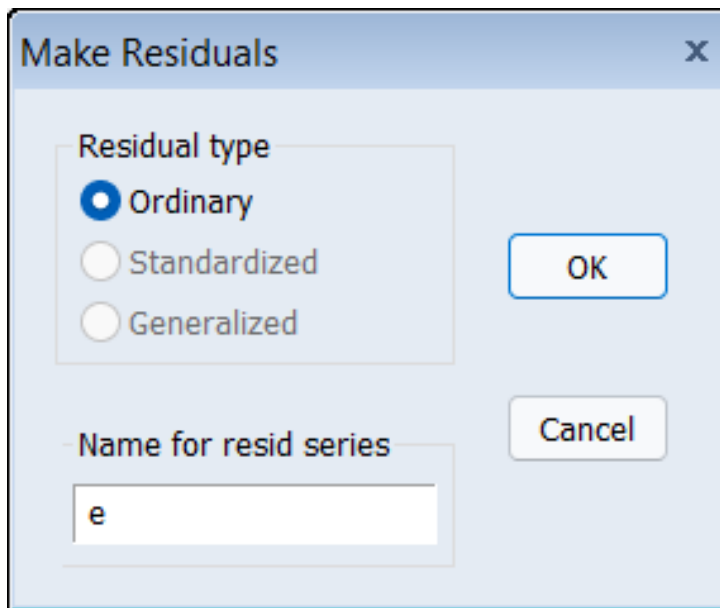
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1299.495	322.8708	-4.024813	0.0003
INCOME	0.319974	0.010502	30.46757	0.0000
R-squared	0.965671	Mean dependent var	5821.629	
Adjusted R-squared	0.964630	S.D. dependent var	7007.005	
S.E. of regression	1317.797	Akaike info criterion	17.26076	
Sum squared resid	57307427	Schwarz criterion	17.34963	
Log likelihood	-300.0632	Hannan-Quinn criter.	17.29144	
F-statistic	928.2730	Durbin-Watson stat	0.190961	
Prob(F-statistic)	0.000000			

第四节 金融实例分析

2. 异方差检验

(1) 图示法

在回归结果窗口，选择“Proc” → “Make Residual Series”，得到残差序列 e 。

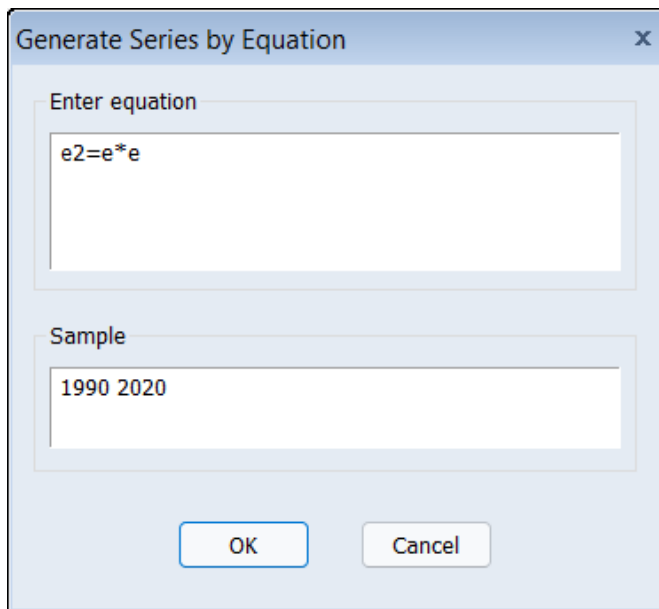


第四节 金融实例分析

2. 异方差检验

(1) 图示法

选择“Qucik” → “Generate Series”，得到残差平方 $e2$ 。



The screenshot shows a dialog box titled "Generate Series by Equation". It has two main input fields: "Enter equation" and "Sample". The "Enter equation" field contains the text "e2=e*e". The "Sample" field contains the text "1990 2020". At the bottom of the dialog box, there are two buttons: "OK" and "Cancel".

第四节 金融实例分析

2. 异方差检验

(1) 图示法

图3-3 残差序列

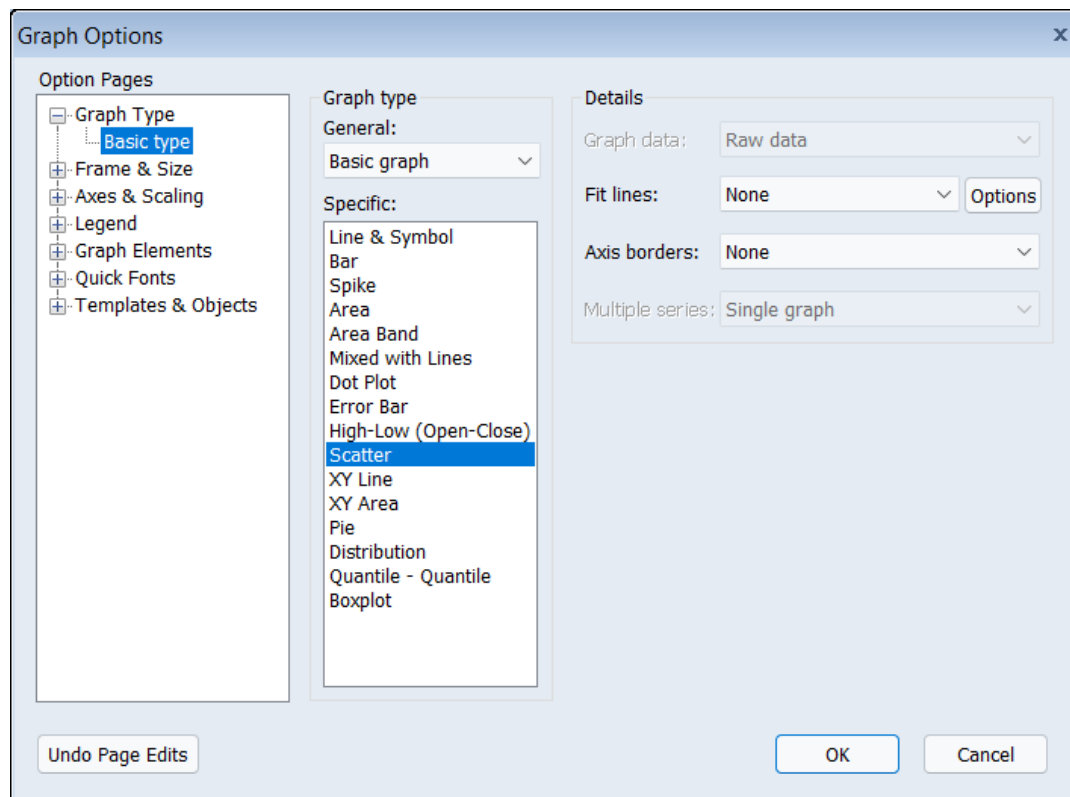
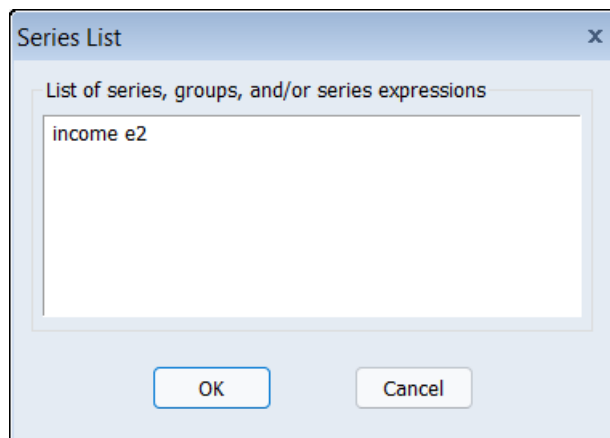
Series: E Workfile: EX3.1::Untitled\				
View	Proc	Object	Properties	Print Name Freeze Default
E				
Last updated: 03/22/22 - 13:38				
Modified: 1985 2019 // eq02.makesresid				
1985	1038.522			
1986	1008.768			
1987	994.6914			
1988	823.1787			
1989	831.2252			
1990	846.9905			
1991	823.0382			
1992	836.6915			
1993	677.9639			
1994	620.8845			
1995	308.6378			
1996	84.82301			

图3-4 残差平方序列

Series: E2 Workfile: EX3.1::Untitled\				
View	Proc	Object	Properties	Print Name Freeze Default
717392.838383698				
Last updated: 03/22/22 - 13:41				
Modified: 1985 2019 // e2=e*e				
1985	1078528.			
1986	1017612.			
1987	989411.0			
1988	677623.2			
1989	690935.3			
1990	717392.8			
1991	677391.9			
1992	700052.7			
1993	459635.1			
1994	385497.6			
1995	95257.30			
1996	7194.943			

第四节 金融实例分析

选择“Qucik” → “Graph”，在“Series List”窗口输入：income e2，点击“OK”；然后在“Graph Options”窗口选择“Scatter”（散点图）。



第四节 金融实例分析

2. 异方差检验

(1) 图示法

从图中大致可以看出，随机误差项存在异方差，随机误差项的平方可能与人均收入存在正相关关系。

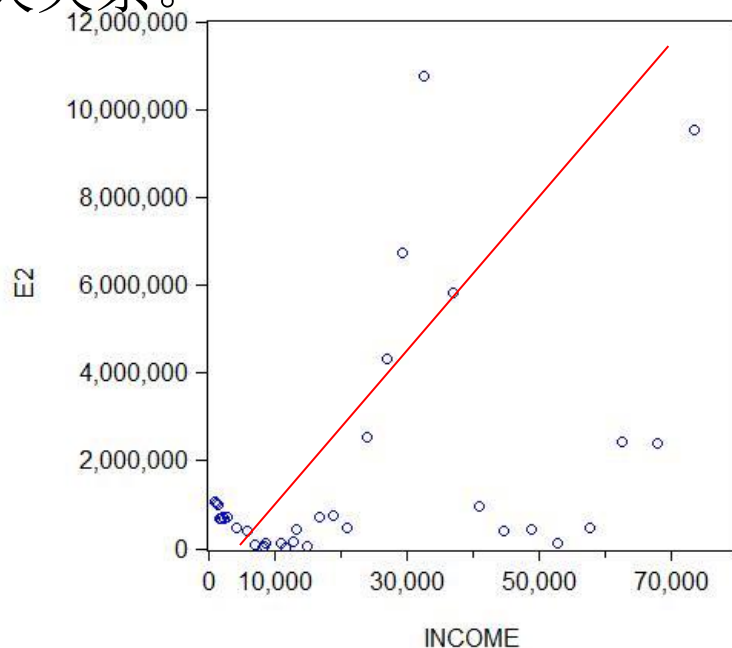


图3-5 残差图

第四节 金融实例分析

2. 异方差检验

(2) Goldfeld-Quandt检验

首先将时间定义为1985-1998年，然后用OLS方法进行估计：

Equation Estimation

Specification Options

Equation specification

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)*X$.

saving c income

Estimation settings

Method: LS - Least Squares (NLS and ARMA)

Sample: 1985 1998

确定 取消

第四节 金融实例分析

2. 异方差检验

(2) Goldfeld-Quandt检验

然后以相同的方法对2006-2019年的样本进行估计。

表3-2 OLS估计结果（1985-1998）

Dependent Variable: SAVING
Method: Least Squares
Date: 03/22/22 Time: 13:47
Sample: 1985 1998
Included observations: 14

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-198.8015	42.00651	-4.732636	0.0005
INCOME	0.217089	0.008390	25.87480	0.0000
R-squared	0.982392	Mean dependent var	698.7857	
Adjusted R-squared	0.980925	S.D. dependent var	641.7663	
S.E. of regression	88.63678	Akaike info criterion	11.93853	
Sum squared resid	94277.74	Schwarz criterion	12.02983	
Log likelihood	-81.56974	Hannan-Quinn criter.	11.93008	
F-statistic	669.5054	Durbin-Watson stat	1.459113	
Prob(F-statistic)	0.000000			

表3-3 OLS估计结果（2006-2019）

Dependent Variable: SAVING
Method: Least Squares
Date: 03/22/22 Time: 13:48
Sample: 2006 2019
Included observations: 14

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-6006.438	734.6499	-8.175919	0.0000
INCOME	0.413687	0.015548	26.60622	0.0000
R-squared	0.983331	Mean dependent var	12347.21	
Adjusted R-squared	0.981942	S.D. dependent var	7035.741	
S.E. of regression	945.4709	Akaike info criterion	16.67281	
Sum squared resid	10726983	Schwarz criterion	16.76410	
Log likelihood	-114.7096	Hannan-Quinn criter.	16.66436	
F-statistic	707.8909	Durbin-Watson stat	0.643093	
Prob(F-statistic)	0.000000			



第四节 金融实例分析

2. 异方差检验

(2) Goldfeld-Quandt检验

1985-1998年OLS估计结果:

$$\text{saving} = -198.8015 + 0.2171\text{income} \quad (1985-1998) \\ (-4.7326) \quad (25.8748)$$

$$R^2 = 0.9824 \quad \sum \hat{\varepsilon}_1^2 = 94277.74$$

2006-2019年OLS方法估计结果:

$$\text{saving} = -6006.438 + 0.4137\text{income} \quad (2006-2019) \\ (-8.1759) \quad (26.6062)$$

$$R^2 = 0.9833 \quad \sum \hat{\varepsilon}_2^2 = 10726983$$

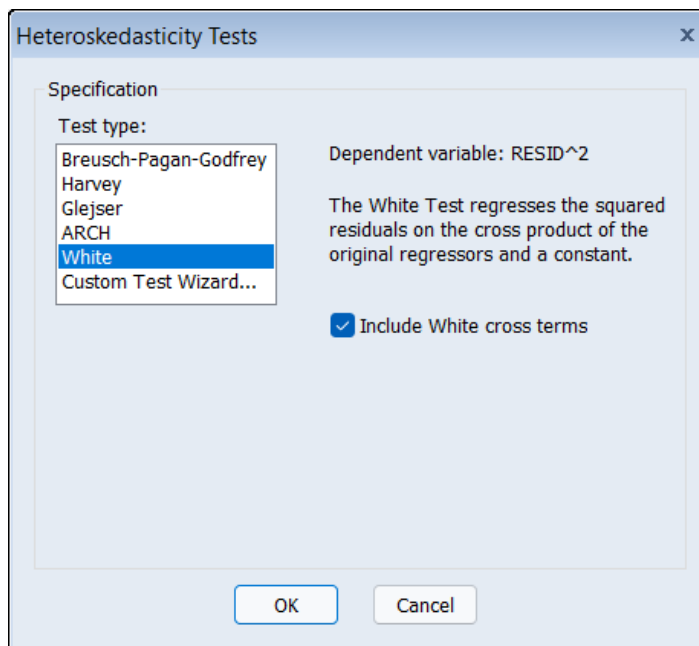
$F = \frac{\sum \hat{\varepsilon}_2^2}{\sum \hat{\varepsilon}_1^2} = 113.78 > F(12,12)_{0.05} = 2.69$ 拒绝 $H_0: \sigma_1^2 = \sigma_2^2$, 表明随机误差项显著存在异方差。

第四节 金融实例分析

2. 异方差检验

(3) White检验

在全体样本的OLS估计结果窗口，选择“View” → “Residual Diagnostics” → “Heteroskedasticity Tests”，选择“Test type”为“White”。



第四节 金融实例分析

(3) White检验

根据表3-4，White检验的统计量 nR^2 对应的p值为0.0298，在5%显著性水平下拒绝原假设。因此，根据White检验，认为模型中存在异方差。

表3-4 White检验结果

Heteroskedasticity Test: White

F-statistic	4.016815	Prob. F(2,32)	0.0278
Obs*R-squared	7.023521	Prob. Chi-Square(2)	0.0298
Scaled explained SS	7.911667	Prob. Chi-Square(2)	0.0191

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 03/22/22 Time: 13:58
Sample: 1985 2019
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	285976.7	784235.5	0.364657	0.7178
INCOME^2	-0.000265	0.001025	-0.258216	0.7979
INCOME	71.96628	69.17370	1.040371	0.3060
R-squared	0.200672	Mean dependent var	1637355.	
Adjusted R-squared	0.150714	S.D. dependent var	2644609.	
S.E. of regression	2437185.	Akaike info criterion	32.33240	
Sum squared resid	1.90E+14	Schwarz criterion	32.46572	
Log likelihood	-562.8170	Hannan-Quinn criter.	32.37842	
F-statistic	4.016815	Durbin-Watson stat	0.694113	
Prob(F-statistic)	0.027772			

第四节 金融实例分析

3. 异方差的修正

(1) 加权最小二乘法

“Qucik” → “Estimate Equation”

采用方差未知的第一
种情况，即选择权重
为 $\sqrt{Income_t}$

Equation Estimation

Specification Options

Coefficient covariance matrix
Estimation default

☒ d.f. Adjustment

Weights
Type: Inverse std. dev.
Weight series: income^0.5
Scaling: EViews default

Coefficient name
C

ARMA options
Starting coefficient values:
OLS/TSLLS
☒ Backcast MA terms

Iteration control
Max Iterations: 500
Convergence: 0.0001
☐ Display settings

Derivatives
Select method to favor:
☒ Accuracy
☐ Speed
☐ Use numeric only

确定 取消

第四节 金融实例分析

表3-6 WLS估计结果

Dependent Variable: SAVING
Method: Least Squares
Date: 03/22/22 Time: 14:20
Sample: 1985 2019
Included observations: 35
Weighting series: INCOME^0.5
Weight type: Inverse standard deviation (EViews default scaling)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3294.775	515.4215	-6.392390	0.0000
INCOME	0.366957	0.010857	33.80029	0.0000
Weighted Statistics				
R-squared	0.971926	Mean dependent var	9334.516	
Adjusted R-squared	0.971075	S.D. dependent var	14005.35	
S.E. of regression	1554.676	Akaike info criterion	17.59137	
Sum squared resid	79761596	Schwarz criterion	17.68024	
Log likelihood	-305.8489	Hannan-Quinn criter.	17.62205	
F-statistic	1142.460	Durbin-Watson stat	0.262203	
Prob(F-statistic)	0.000000	Weighted mean dep.	12289.42	
Unweighted Statistics				
R-squared	0.925942	Mean dependent var	5821.629	
Adjusted R-squared	0.923698	S.D. dependent var	7007.005	
S.E. of regression	1935.536	Sum squared resid	1.24E+08	
Durbin-Watson stat	0.077427			

3. 异方差的修正 (1) 加权最小二乘法

第四节 金融实例分析

(1) 加权最小二乘法

对加权最小二乘法的估计得到的残差项进行White检验，检验统计量 nR^2 对应的p值为0.1135，在5%显著性水平下不拒绝原假设。因此认为模型中不存在异方差。

表3-7 White检验结果

Heteroskedasticity Test: White

F-statistic	2.121234	Prob. F(3,31)	0.1177
Obs*R-squared	5.961122	Prob. Chi-Square(3)	0.1135
Scaled explained SS	5.618124	Prob. Chi-Square(3)	0.1317

Test Equation:

Dependent Variable: WGT_RESID^2

Method: Least Squares

Date: 03/22/22 Time: 14:21

Sample: 1985 2019

Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-371758.6	1284312.	-0.289461	0.7742
INCOME^2*WGT^2	0.001566	0.001181	1.325837	0.1946
INCOME*WGT^2	-178.3915	123.9595	-1.439111	0.1601
WGT^2	6084391.	3466897.	1.754996	0.0891
R-squared	0.170318	Mean dependent var		2278903.
Adjusted R-squared	0.090026	S.D. dependent var		3366830.
S.E. of regression	3211705.	Akaike info criterion		32.90971
Sum squared resid	3.20E+14	Schwarz criterion		33.08747
Log likelihood	-571.9200	Hannan-Quinn criter.		32.97107
F-statistic	2.121234	Durbin-Watson stat		0.891006
Prob(F-statistic)	0.117660			

第四节 金融实例分析

3. 异方差的修正

(2) 对数变换法

“Quick” → “Estimate Equation”，对模型中的所有变量取对数。

Equation Estimation

Specification Options

Equation specification

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)*X$.

log(saving) c log(income)

Estimation settings

Method: LS - Least Squares (NLS and ARMA)

Sample: 1985 2019

确定 取消



第四节 金融实例分析

3. 异方差的修正 (2) 对数变换法

表3-8 对数变换估计结果

Dependent Variable: LOG(SAVING)
Method: Least Squares
Date: 03/22/22 Time: 14:25
Sample: 1985 2019
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.569937	0.285484	-16.00767	0.0000
LOG(INCOME)	1.307484	0.030130	43.39411	0.0000
R-squared	0.982777	Mean dependent var	7.708128	
Adjusted R-squared	0.982255	S.D. dependent var	1.687896	
S.E. of regression	0.224844	Akaike info criterion	-0.091372	
Sum squared resid	1.668313	Schwarz criterion	-0.002495	
Log likelihood	3.599013	Hannan-Quinn criter.	-0.060692	
F-statistic	1883.049	Durbin-Watson stat	1.089167	
Prob(F-statistic)	0.000000			

第四节 金融实例分析

(2) 对数变换法

对对数变换法的估计得到的残差项进行White检验，检验统计量 nR^2 对应的p值为0.1866，在5%显著性水平下不拒绝原假设。因此认为模型中不存在异方差。

表3-9 White检验结果

Heteroskedasticity Test: White

F-statistic	1.697551	Prob. F(2,32)	0.1992
Obs*R-squared	3.357205	Prob. Chi-Square(2)	0.1866
Scaled explained SS	9.757855	Prob. Chi-Square(2)	0.0076

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 03/22/22 Time: 14:26
Sample: 1985 2019
Included observations: 35

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.675054	1.171749	0.576108	0.5686
LOG(INCOME)^2	0.004264	0.014147	0.301421	0.7650
LOG(INCOME)	-0.107577	0.259454	-0.414626	0.6812

R-squared	0.095920	Mean dependent var	0.047666
Adjusted R-squared	0.039415	S.D. dependent var	0.123669
S.E. of regression	0.121207	Akaike info criterion	-1.300810
Sum squared resid	0.470120	Schwarz criterion	-1.167494
Log likelihood	25.76417	Hannan-Quinn criter.	-1.254789
F-statistic	1.697551	Durbin-Watson stat	2.149691
Prob(F-statistic)	0.199209		



第三章 异方差和自相关

第五节 自相关的概念和产生原因

一、滞后值与自相关的概念

二、自相关产生的原因



一、滞后值与自相关的概念

➤ 滞后值：一个变量在一段时间前的取值

例： y_t 滞后一期的取值为 y_{t-1}

➤ y 的一阶差分： $\Delta y_t = y_t - y_{t-1}$

➤ 序列相关/自相关：

在回归模型 $Y_i = \alpha + \beta X_i + \varepsilon$ 中，

$\text{cov}(\varepsilon_i, \varepsilon_j) \neq 0, i \neq j$ ，随机误差项的取值与它的前一期或前几期的取值（滞后值）有关。

↓违反

经典线性回归模型（CLRM）的基本假设第三条： $\text{cov}(\varepsilon_i, \varepsilon_j) = E(\varepsilon_i, \varepsilon_j) = 0, i \neq j, i, j = 1, 2, \dots, n$



一、滞后值与自相关的概念

随机误差项 ε 的自相关现象是经常存在的。在模型的研究中，不可能将所有的影响因素都包括进方差，一些非重要因素会被归入误差项，而这些因素往往有时间趋势，因此使误差项体现了时间先后上的某种相关性。

自相关 { 正相关：对 $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ ，当 $\varepsilon_t > 0$ 时，随后的若干随机项 $\varepsilon_{t+1}, \varepsilon_{t+2} \dots$ 都有大于0的倾向
负相关：两个相继的随机项 $\varepsilon_{t+1}, \varepsilon_{t+2}$ 具有正负号相反的倾向

一、滞后值与自相关的概念

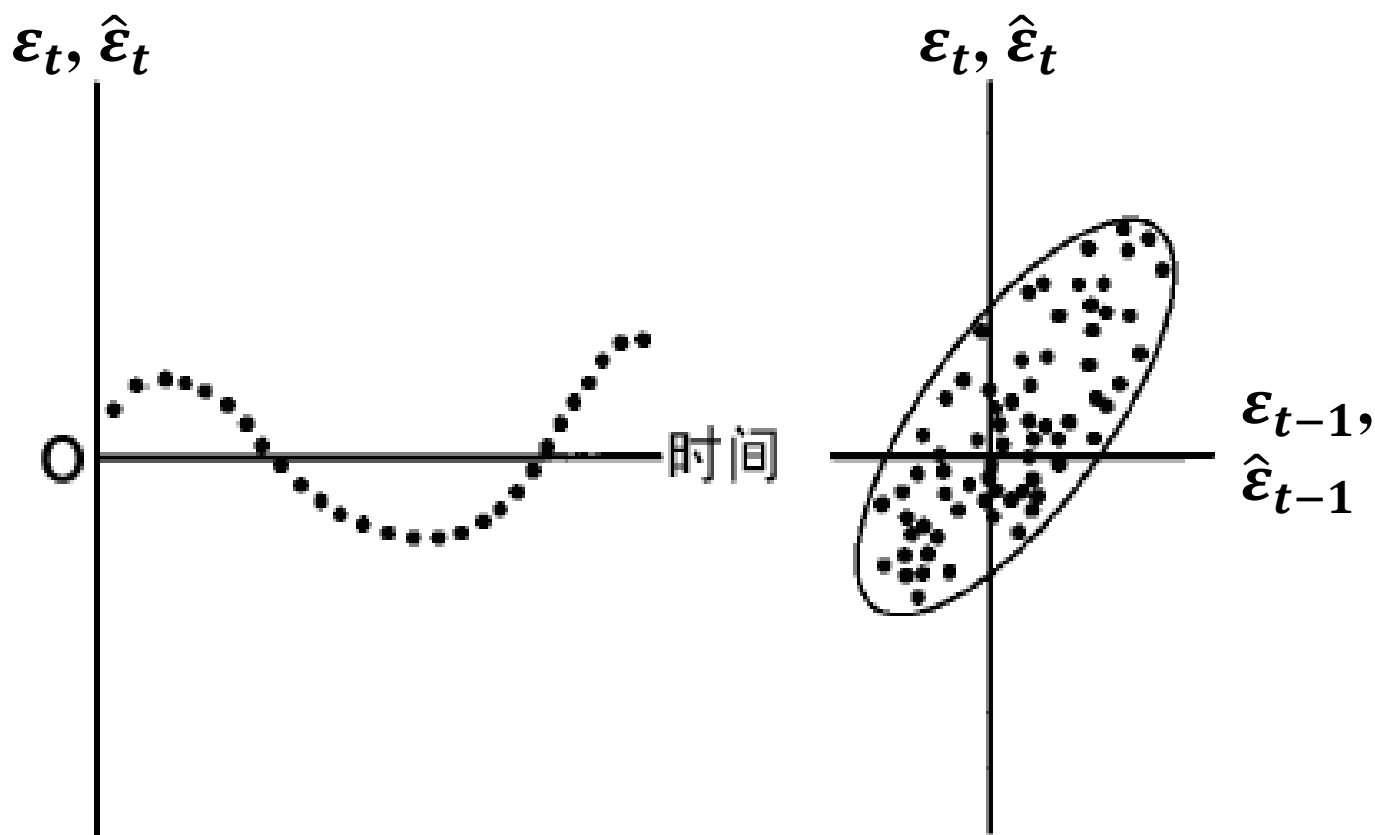


图3-6 (a) 正自相关

一、滞后值与自相关的概念

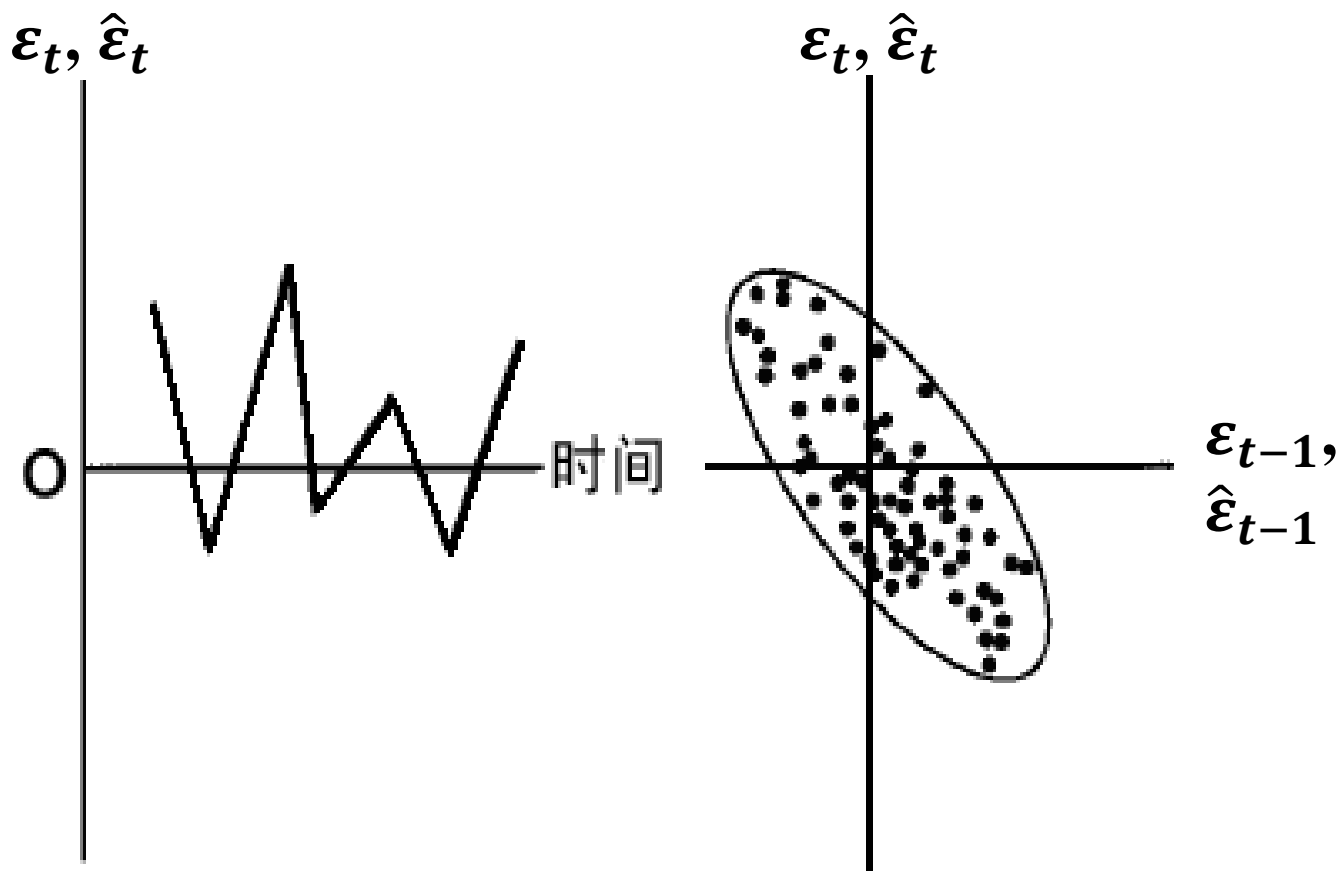


图3-6 (b) 负自相关

一、滞后值与自相关的概念

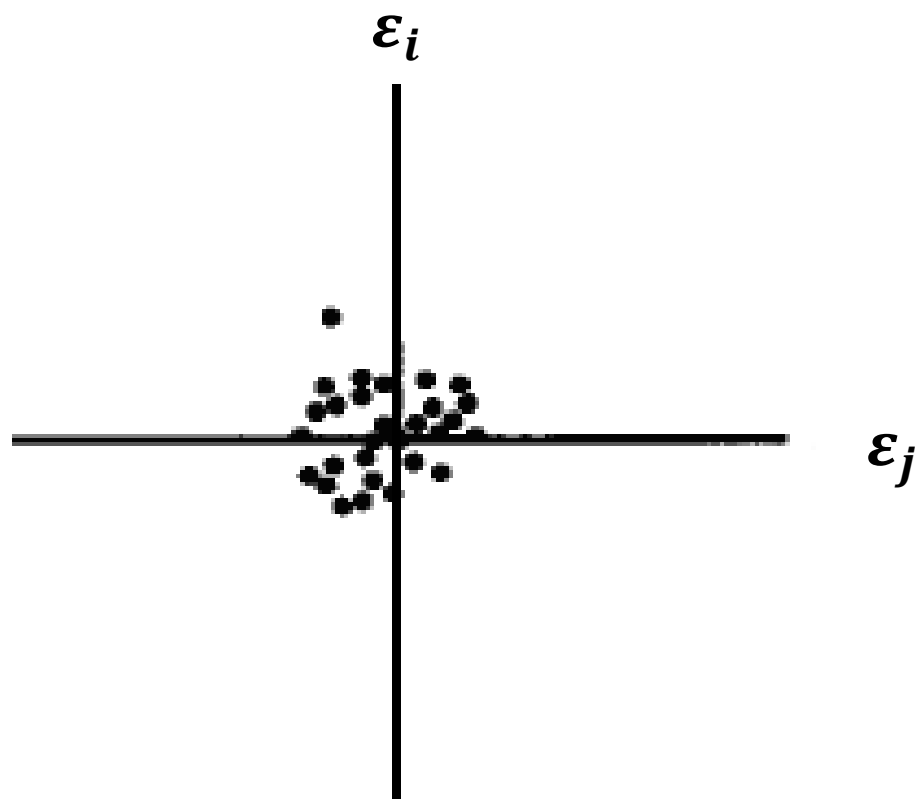


图3-6 (c) 无自相关



二、自相关产生的原因

- (1) 经济数据的固有的惯性 (**inertia**) 带来的相关
 - 时间序列呈现出商业循环（在经济活动中重复发生或自我维持波动）
 - 如**1929**年资本主义经济大危机后，经济开始恢复，多数经济指标序列持续向上移动。
- (2) 模型设定误差带来的相关
 - 本应包括在模型中的重要变量未包括进模型中
 - 选择了错误的函数形式
- (3) 数据的加工带来的相关
 - 对原始数据进行加工，如利用股票价格数据计算收益率
 - 用内插或修匀的方法补充缺少的数据和修正偏差太大的数据



第三章 异方差和自相关

第六节 自相关的度量和后果

一、自相关的度量

二、出现自相关的后果



一、自相关的度量

一阶自相关： ε_t 的取值仅与前期 ε_{t-1} 有关，即 $\varepsilon_t = f(\varepsilon_{t-1})$

- 对于一般经济现象而言，两个随机项在时间上相隔越远，前者对后者的影响越小。如果存在自相关的话，**最强的自相关应该是一阶自相关**。
- 这里只讨论一阶自相关，并且假定这是一种**线性自相关**，具有一阶线性自回归**AR(1)**的形式：

$$\begin{cases} \varepsilon_t = \rho\varepsilon_{t-1} + v_t \\ -1 \leq \rho \leq 1 \end{cases} \quad (3.25)$$

其中 ρ 为常数，称为自相关系数。 v_t 是一个新随机项，满足经典回归的全部假定：

$$\begin{cases} v_t \sim N(0, \sigma^2) \\ cov(v_t, v_s) = 0, t \neq s \end{cases} \quad (3.26)$$



一、自相关的度量

使用OLS法估计 ρ :

$$\hat{\rho} = \frac{\sum_{t=2}^n \varepsilon_t \varepsilon_{t-1}}{\sum_{t=2}^n \varepsilon_t^2} \quad (3.27)$$

$$\rho \left\{ \begin{array}{l} > 0, \text{ 正相关} \\ = 0, \text{ 不存在相关} \\ < 0, \text{ 负相关} \end{array} \right.$$



二、出现自相关的后果

- (1) 最小二乘估计量仍然是线性的和无偏的，但却不是有效的。
- (2) OLS估计量的方差是有偏的。

因此，在随机项存在自相关的情况下，t检验失效，同样对F检验也有类似的结果。



第三章 异方差和自相关

第七节 自相关的检验和修正

一、自相关的检验方法

二、自相关的修正方法

一、自相关的检验方法

（一）图示法

以回归残差 $\hat{\varepsilon}_t$ 估计随机项 ε_t ，通过观察残差是否存在自相关来判断随机项是否存在自相关。

（1）按时间顺序绘制残差图

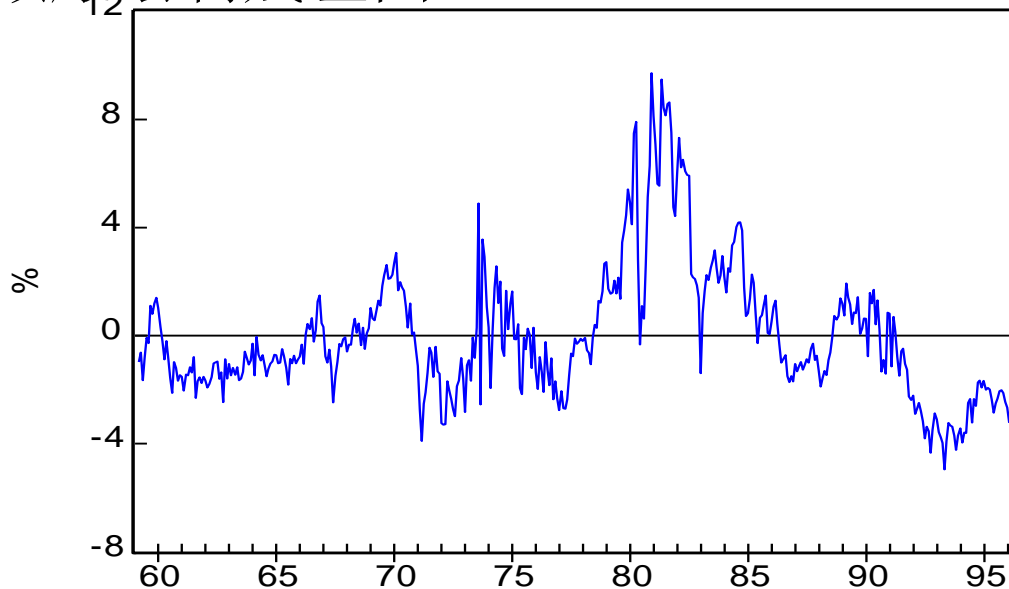


图3-7 利率残差

一、自相关的检验方法

(一) 图示法

(2) 按时间顺序绘制 $\hat{\varepsilon}_t$ 、 $\hat{\varepsilon}_{t-1}$ 散点图

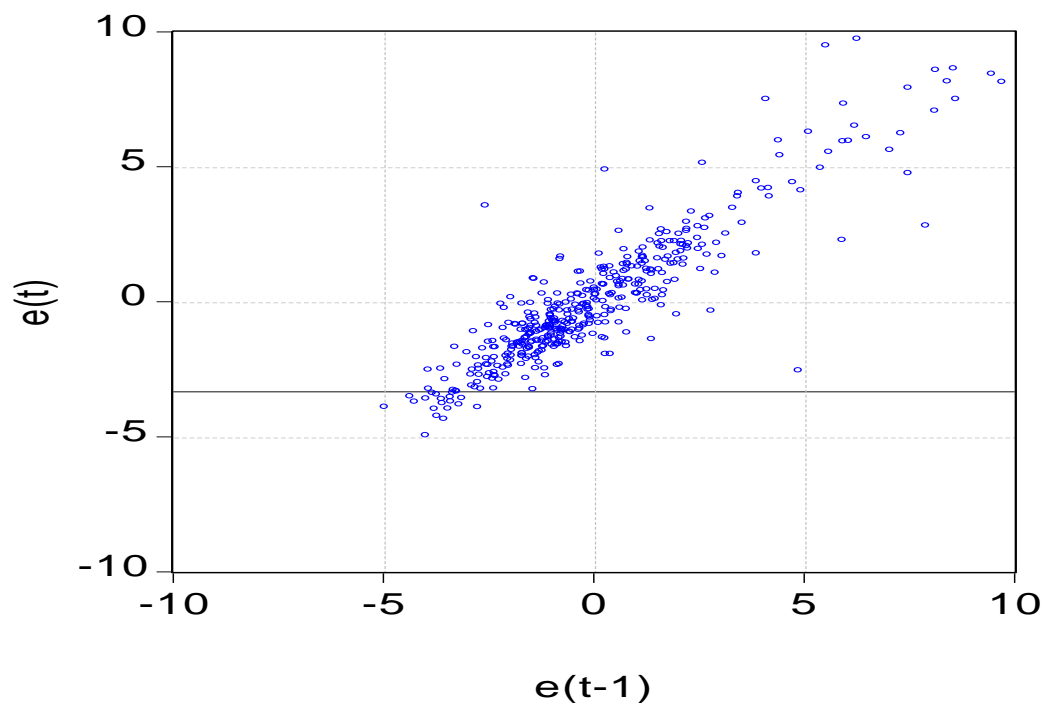


图3-8 利率残差 $\hat{\varepsilon}_t$ 、 $\hat{\varepsilon}_{t-1}$ 散点图



一、自相关的检验方法

（二）解析法

（1）DW（Durbin-Watson）检验

基本思想：对一阶自相关：
$$\begin{cases} \varepsilon_t = \rho\varepsilon_{t-1} + v_t \\ -1 \leq \rho \leq 1 \end{cases}$$

当 $\rho = 0$ 时，不具有自相关，当 $\rho \neq 0$ 时，具有自相关。

构造统计量： $d = 2 - 2 \sum_{t=2}^n \hat{\varepsilon}_t \hat{\varepsilon}_{t-1} / \sum_{t=1}^n \hat{\varepsilon}_t^2 \approx 2(1 - \hat{\rho})$

表3-10 d 与 $\hat{\rho}$ 的近似关系

$\hat{\rho}$ 值	d 值（近似）
-1（完全负相关）	4
0（没有自相关）	2
1（完全正相关）	0

一、自相关的检验方法

(二) 解析法

(1) DW (Durbin-Watson) 检验

Durbin-Watson证明了 d 的实际分布介于两个极限分布之间：

下极限分布： $(d_L, 4 - d_U)$ 上极限分布： $(d_U, 4 - d_L)$

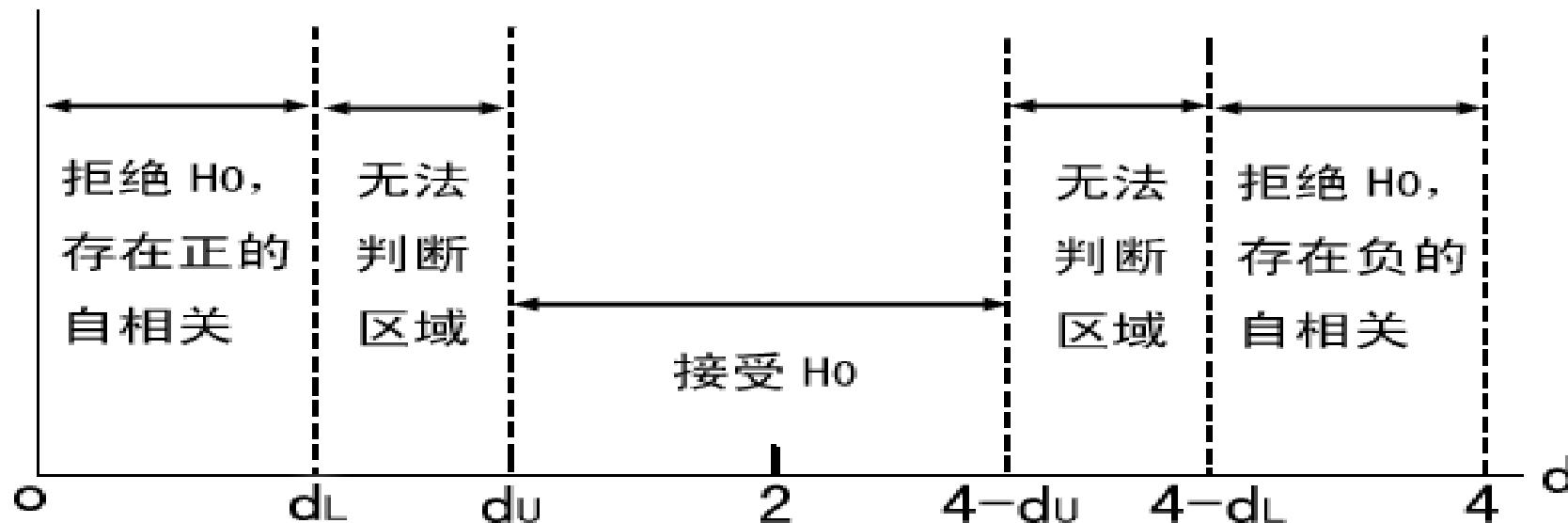


图3-9 Durbin-Watson d 统计量



一、自相关的检验方法

（二）解析法

（1）DW（Durbin-Watson）检验

检验步骤：

1. 建立假设： $H_0: \rho = 0; H_1: \rho \neq 0$
2. 进行OLS回归并获得残差；
3. 计算 d 值，大多数计算软件已能够实现。比如：**Eviews**软件就直接可以获得；
4. 给定样本容量及解释变量的个数，从DW表中查到临界值 d_L 和 d_U ；
5. 将 d 的现实值与临界值进行比较。



一、自相关的检验方法

（二）解析法

（1）DW（Durbin-Watson）检验

局限性：

1. DW检验不适合用于自回归模型。
2. D-W检验只适用于一阶线性自相关，对高阶自相关皆不适用。
3. d 统计量无法用来判定那些通过原点的回归模型的自相关问题。
4. 利用D-W检验检验自相关时，一般要求样本容量至少为**15**，否则很难对自相关的存在性做明确的结论。



一、自相关的检验方法

(二) 解析法

(2) 杜宾—h (Durbin-h) 统计量

- 适用情况：解释变量中包含有因变量的滞后值
- 基于h统计量的渐近检验：

$$h = \hat{\rho} \sqrt{\frac{T}{1 - T[\text{var}(\hat{\beta})]}} \quad (3.29)$$

其中， $\hat{\rho} = \sum_{t=2}^n \hat{\varepsilon}_t \hat{\varepsilon}_{t-1} / \sum_{t=2}^n \hat{\varepsilon}_{t-1}^2$ ， $\hat{\varepsilon}_t$ 为残差， $\text{var}(\hat{\beta})$ 是因变量滞后项的系数的OLS估计量的方差。

在没有自相关的原假设之下，h统计量是**渐近正态**的，其均值为0，方差为1。当检验一阶自回归的误差时，即使X包含有多个因变量的滞后值，统计量检验仍然有效。



一、自相关的检验方法

(二) 解析法

(3) Breusch-Godfrey检验

- 适用情况：序列可能存在高阶自相关；解释变量中包含有因变量的滞后值
- 缺陷：无法确定滞后阶数 r 的值。
- **BG**检验法假定误差项是由如下的阶自回归过程产生的：

$$\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \rho_3 \varepsilon_{t-3} + \dots + \rho_r \varepsilon_{t-r} + v_t$$

其中， $v_t \sim N(0, \sigma_v^2)$ 。

$$H_0: \rho_1 = \rho_2 = \rho_3 = \dots = \rho_r = 0$$



一、自相关的检验方法

(二) 解析法

(3) Breusch-Godfrey检验

➤ 检验步骤:

1. 用最小二乘法估计回归模型并得到残差 $\hat{\varepsilon}_t$ 。

2. 将 $\hat{\varepsilon}_t$ 对第一步中的所有解释变量及 $\hat{\varepsilon}_t$ 的 r 个滞后值
($\hat{\varepsilon}_{t-1}, \hat{\varepsilon}_{t-2}, \dots, \hat{\varepsilon}_{t-r}$) 进行回归, 并取得 R^2 值。

3. 建立检验统计量 $(T-r)R^2$,

在大样本的条件下, $(T-r)R^2 \sim \chi^2(r)$ 。

若 $(T-r)R^2$ 大于临界值, 则拒绝不存在自相关的零假设,
反之则不能拒绝。



二、自相关的修正方法

(一) ρ 已知——广义差分法

回归模型:

$$Y_t = b_0 + b_1X_{1t} + b_2X_{2t} + b_3X_{3t} + \varepsilon_t \quad (3.31)$$

假定残差项存在一阶自回归:

$\varepsilon_t = \rho\varepsilon_{t-1} + v_t$, v_t 满足经典回归的全部假定。

若经过DW检验发现随机项具有正的自相关, 已知 $d \approx 2(1 - \hat{\rho})$,

则 $\hat{\rho} \approx 1 - 1/2d$

将式 (3.31) 滞后一期并乘以 $\hat{\rho}$:

$$\hat{\rho}Y_{t-1} = \hat{\rho}b_0 + \hat{\rho}b_1X_{1,t-1} + \hat{\rho}b_2X_{2,t-1} + \hat{\rho}b_3X_{3,t-1} + \hat{\rho}\varepsilon_{t-1} \quad (3.32)$$

用式 (3.31) 减去 (3.32):

$$Y_t - \hat{\rho}Y_{t-1} = (1 - \hat{\rho})b_0 + b_1(X_{1t} - \hat{\rho}X_{1,t-1}) + b_2(X_{2t} - \hat{\rho}X_{2,t-1}) + b_3(X_{3t} - \hat{\rho}X_{3,t-1}) + \varepsilon_t - \hat{\rho}\varepsilon_{t-1} \quad (3.33)$$



二、自相关的修正方法

(一) ρ 已知——广义差分法

令

$$dY_t = Y_t - \hat{\rho}Y_{t-1}$$

$$dX_{1t} = X_{1t} - \hat{\rho}X_{1,t-1}$$

$$dX_{2t} = X_{2t} - \hat{\rho}X_{2,t-1}$$

$$dX_{3t} = X_{3t} - \hat{\rho}X_{3,t-1}$$

$$a = (1 - \hat{\rho})b_0$$

(3.34)

广义差分变换

广义差分模型

$$dY_t = a + b_1dX_{1t} + b_2dX_{2t} + b_3dX_{3t} + v_t \quad (3.35)$$

其中， $v_t = \varepsilon_t - \hat{\rho}\varepsilon_{t-1}$ ，满足经典回归的全部假定，可以对模型 (3.35) 进行OLS估计。



二、自相关的修正方法

(一) ρ 已知——广义差分法

经过广义差分变换后，数据损失一个观测值，为避免该损失，**K. R. Kadiyala**提出对第一个观测值作如下变换：

$$\begin{aligned}dx_1 &= \sqrt{1 - \hat{\rho}^2} x_1 \\dy_1 &= \sqrt{1 - \hat{\rho}^2} y_1\end{aligned}$$



二、自相关的修正方法

（一） ρ 已知——广义差分法

自相关的具体解决方案：**Cochrane-Orcutt**（简称**CO法**）

1. 根据原始数据，利用**OLS**法对原模型进行参数估计，得出拟合值及残差。

2. 假设残差存在一阶自相关，利用**OLS**方法对 $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$ 进行估计，得到 $\hat{\rho}$ ，若 $\hat{\rho}$ 满足精度，则停止；否则，转入下一步。

3. 进行广义差分变换： $y_t^* = y_t - \hat{\rho}y_{t-1}$ ， $x_{it}^* = x_{it} - \hat{\rho}x_{i,t-1}$ ， $\varepsilon_t^* = \varepsilon_t - \hat{\rho}\varepsilon_{t-1}$ ，（ $i = 1, 2, \dots, k; t = 2, 3, \dots, T$ ），求得模型 $y_t^* = \beta_1(1 - \hat{\rho}) + \beta_2x_{2t}^* + \dots + \beta_kx_{kt}^* + v_t$ ，对该模型重新估计得到原方程的参数估计值。将这些参数估计代入原方程得到新的回归残差，重复步骤2。

迭代停止条件： $\hat{\rho}$ 前后两次估计之差小于**0.01**或**0.05**；或迭代超过一定（如**10, 20...**）次数。



二、自相关的修正方法

(一) ρ 未知——杜宾两步法

第一步:

$$Y_t - \rho Y_{t-1} = (1 - \rho)b_0 + b_1(X_{1t} - \rho X_{1,t-1}) + b_2(X_{2t} - \rho X_{2,t-1}) + b_3(X_{3t} - \rho X_{3,t-1}) + \varepsilon_t - \rho \varepsilon_{t-1}$$

变换

$$Y_t = (1 - \rho)b_0 + \rho Y_{t-1} + b_1 X_{1t} - \rho b_1 X_{1,t-1} + b_2 X_{2t} - \rho b_2 X_{2,t-1} + b_3 X_{3t} - \rho b_3 X_{3,t-1} + v_t$$

对变换后的模型进行OLS估计, 得到 Y_{t-1} 的系数估计值 $\hat{\rho}$ 。



二、自相关的修正方法

(一) ρ 未知——杜宾两步法

第二步:

用估计值 $\hat{\rho}$ 对原始数据进行差分变换, 得到:

$$dY_t = a + b_1 dX_{1t} + b_2 dX_{2t} + b_3 dX_{3t} + v_t$$

对上式进行OLS估计, 得到:

$$\widehat{dY}_t = \hat{a} + \hat{b}_1 dX_{1t} + \hat{b}_2 dX_{2t} + \hat{b}_3 dX_{3t}$$

用杜宾两步法修正的结果为:

$$\hat{Y}_t = \hat{a}/(1 - \hat{\rho}) + \hat{b}_1 X_{1t} + \hat{b}_2 X_{2t} + \hat{b}_3 X_{3t}$$



例3.2 利率的变化

我们将用工业生产指数（ IP ）、货币供应量增长率（ $GM2$ ），以及通货膨胀率（ $GCPI$ ）来解释国债利率（ R ）的变化。

IP 表示联邦储备委员会的工业生产指数；

$GM2_t = \frac{M2_t - M2_{t-1}}{M2_{t-1}}$ ，其中 $M2$ 表示名义货币供给，以10亿美元为单位；

$GCPI_t = \frac{CPI_t - CPI_{t-1}}{CPI_{t-1}}$ ， CPI 表示消费者物价指数；

R 表示3个月期美国国债利率。

采用2002年1月—2021年12月的月度数据，建立如下回归模型：

$$R_t = a + b_1 IP_t + b_2 GM2_t + b_3 GCPI_t + \varepsilon_t$$

例3.2 利率的变化

1. 自相关的检验

对原始回归模型进行最小二乘回归

Equation Estimation

Specification Options

Equation specification

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)*X$.

r c ip gm2 gcpi(-1)

Estimation settings

Method: LS - Least Squares (NLS and ARMA)

Sample: 2002M01 2021M12

确定 取消

例3.2 利率的变化

1. 自相关的检验-DW检验

回归模型的d值为0.036049，样本中有239个观测值。而DW表中样本容量最大为 $n=200$ 。在1%显著性水平下， $n=200$ ， $k=3$ 时， $d_L = 1.643$ ， $d < d_L$ ，因此，我们在1%显著性水平下拒绝原假设，随机项有很强的正相关。

表3-11 原始模型的最小二乘回归结果

Dependent Variable: R

Method: Least Squares

Date: 02/18/22 Time: 19:37

Sample (adjusted): 2002M02 2021M12

Included observations: 239 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5.560142	1.931382	-2.878841	0.0044
IP	0.070560	0.019704	3.580967	0.0004
GM2	-26.63531	15.07317	-1.767067	0.0785
GCPI(-1)	27.68259	24.43360	1.132972	0.2584
R-squared	0.083821	Mean dependent var		1.197711
Adjusted R-squared	0.072125	S.D. dependent var		1.467440
S.E. of regression	1.413530	Akaike info criterion		3.546652
Sum squared resid	469.5459	Schwarz criterion		3.604836
Log likelihood	-419.8250	Hannan-Quinn criter		3.570099
F-statistic	7.166732	Durbin-Watson stat		0.036049
Prob(F-statistic)	0.000126			

例3.2 利率的变化

1. 自相关的检验-按时间顺序绘制残差图

也可以通过图示法观察残差是否存在自相关。在回归结果的界面选择

“View” → “Actual, Fitted, Residual” → “Residual Graph”。从图中可以看出，回归残差高度正相关。某一时期的残差若为正（负），则下一时期的残差也很有可能为正（负）。

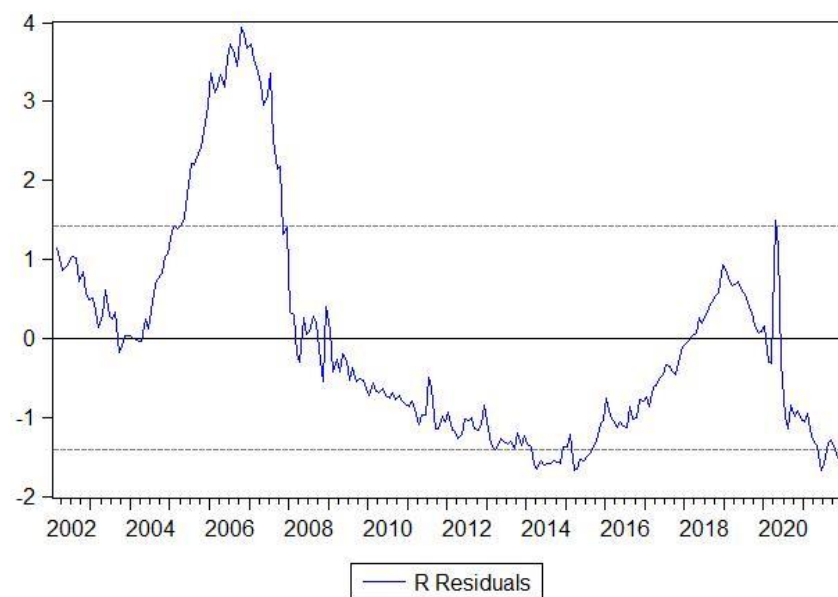


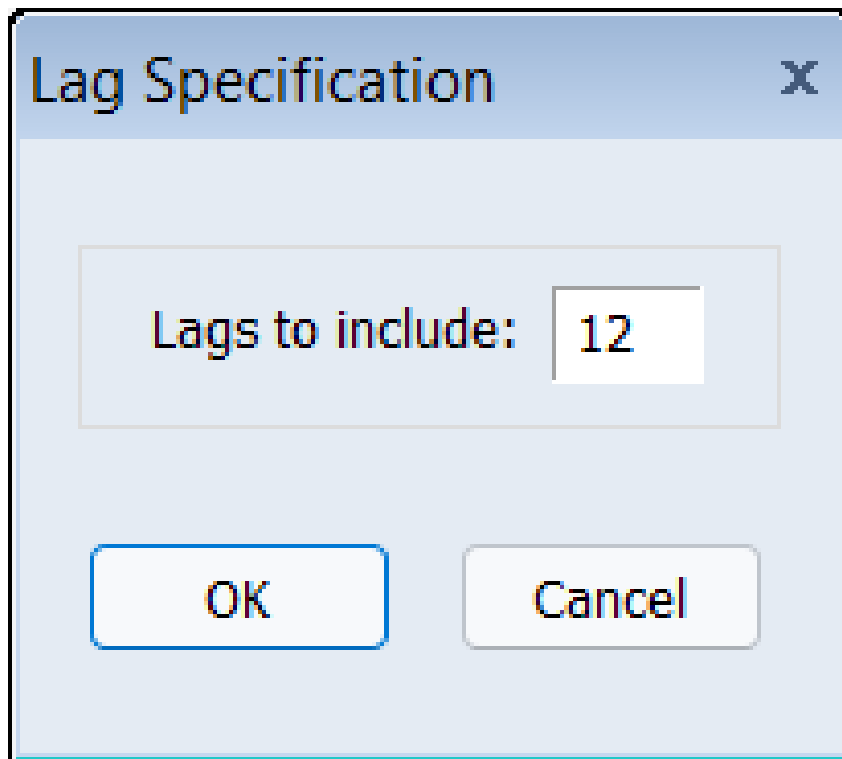
图3-9 利率残差



例3.2 利率的变化

1. 自相关的检验 - Breusch-Godfrey检验

在OLS估计结果窗口，点击“View” → “Residual Diagnostics” → “Series Correlation LM Test”。本例中使用月度数据，因此选择滞后阶数为12阶。



1. 自相关的检验 - Breusch-Godfrey检验

Breusch-Godfrey检验建立的统计量为 $(T - r)R^2$ ，本例中 $(T - r)R^2$ 的值为231.7272，对应的p值为0.00，因此拒绝不存在自相关的零假设，认为残差序列存在自相关。

表3-12 Breusch-Godfrey检验结果

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	592.1095	Prob. F(12,223)	0.0000
Obs*R-squared	231.7272	Prob. Chi-Square(12)	0.0000

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 02/25/22 Time: 10:54

Sample: 2002M02 2021M12

Included observations: 239

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.208511	0.349685	0.596283	0.5516
IP	-0.002719	0.003569	-0.762033	0.4468
GM2	12.34198	2.793368	4.418315	0.0000
GCPI(-1)	-9.666743	4.591085	-2.105546	0.0364
RESID(-1)	0.901292	0.065219	13.81957	0.0000
RESID(-2)	-0.125501	0.090679	-1.384008	0.1677
RESID(-3)	0.281749	0.089523	3.147220	0.0019
RESID(-4)	-0.058288	0.091458	-0.637325	0.5246
RESID(-5)	0.177721	0.091330	1.945929	0.0529
RESID(-6)	-0.170753	0.091682	-1.862455	0.0639
RESID(-7)	0.142899	0.091640	1.559344	0.1203
RESID(-8)	-0.120451	0.091438	-1.317294	0.1891
RESID(-9)	0.065318	0.092028	0.709765	0.4786
RESID(-10)	-0.035680	0.089860	-0.397056	0.6917
RESID(-11)	-0.002979	0.088959	-0.033492	0.9733
RESID(-12)	-0.083611	0.063963	-1.307177	0.1925
R-squared	0.969570	Mean dependent var	-1.92E-15	
Adjusted R-squared	0.967523	S.D. dependent var	1.404593	
S.E. of regression	0.253126	Akaike info criterion	0.154742	
Sum squared resid	14.28824	Schwarz criterion	0.387476	
Log likelihood	-2.491638	Hannan-Quinn criter.	0.248527	
F-statistic	473.6876	Durbin-Watson stat	1.608610	
Prob(F-statistic)	0.000000			



例3.2 利率的变化

2. 自相关的修正——广义差分法

假定残差项存在一阶自回归：

$\varepsilon_t = \rho\varepsilon_{t-1} + v_t$ ， v_t 满足经典回归的全部假定。

经过DW检验，已知 $d = 0.036049$ ，则 $\hat{\rho} \approx 1 - 1/2DW = 0.9819755$

对所有变量进行广义差分变换：

$$\left\{ \begin{array}{l} dR_t = R_t - \hat{\rho}R_{t-1} \\ dIP_t = IP_t - \hat{\rho}IP_{t-1} \\ dGM2_t = GM2_t - \hat{\rho}GM2_{t-1} \\ dGCPI_t = GCPI_t - \hat{\rho}GCPI_{t-1} \\ a = (1 - \hat{\rho})b_0 \end{array} \right.$$

得到广义差分模型：

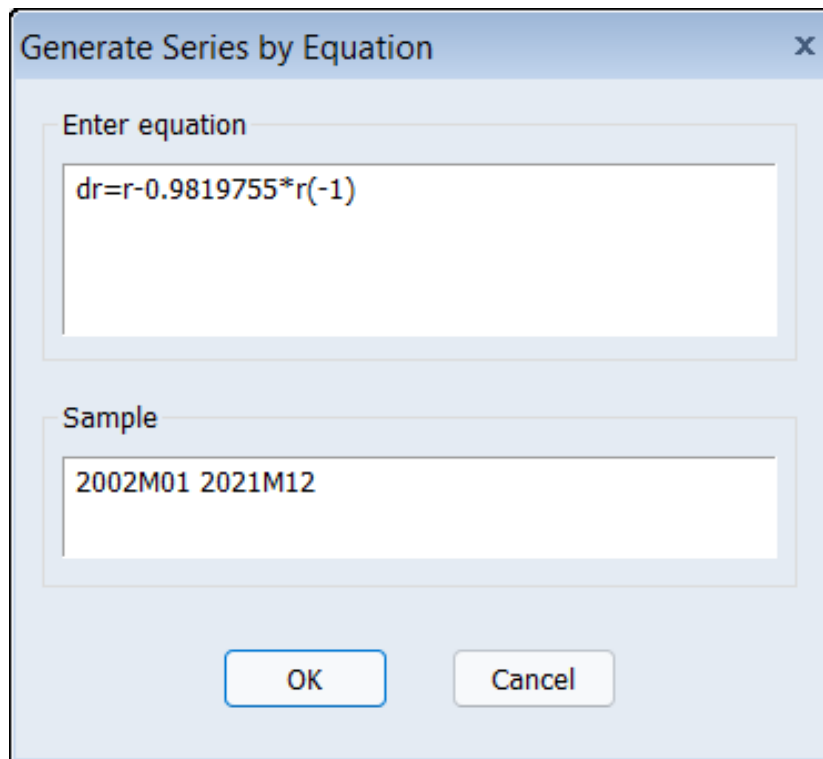
$$dR_t = a + b_1dIP_t + b_2dGM2_t + b_3dGCPI_{t-1} + v_t$$



例3.2 利率的变化

2. 自相关的修正——广义差分法

以变量 R 为例，在Eviews进行广义差分变换：“Quick” → “Generate Series”。



Generate Series by Equation

Enter equation

$dr=r-0.9819755*r(-1)$

Sample

2002M01 2021M12

OK Cancel

可以看到生成的变量 dR 损失了第一个观测值，采用Kadiyala提出的方法，对第一个观测值作如下变换：

$$dR_1 = \sqrt{1 - \hat{\rho}^2} R_1 = \sqrt{1 - 0.9819755^2} \times 1.777 = 0.335867655$$

在变量dR的窗口点击“Edit+/-”，将数值0.981317874填入第一个观测值。

[illegible]

例3.2 利率的变化

2. 自相关的修正——广义差分法

用同样的方法对所有变量进行广义差分变换后，采用OLS法估计广义差分模型：“Quick” → “Estimate Equation”。

Equation Estimation

Specification Options

Equation specification

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)*X$.

dr c dip dgm2 dgcpi(-1)

Estimation settings

Method: LS - Least Squares (NLS and ARMA)

Sample: 2002M01 2021M12

确定 取消

2. 自相关的修正——广义差分法

- 修正后模型的拟合优度反而下降;
- 修正后DW检验的 d 统计量值为1.792711, $d_U = 1.704$, $d_U < d < 4 - d_U$, 因此接受原假设, 即认为残差项不具有一阶自相关;
- 修正后的货币供应量增长率 ($GM2$) 在5%的水平下显著为负, 相比于原来的估计结果更加显著;
- 修正后的通货膨胀率在5%水平下显著为正, 符合理论预期。

表3-13 原始模型的最小二乘回归结果

Dependent Variable: R
Method: Least Squares
Date: 02/18/22 Time: 19:37
Sample (adjusted): 2002M02 2021M12
Included observations: 239 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-5.560142	1.931382	-2.878841	0.0044
IP	0.070560	0.019704	3.580967	0.0004
GM2	-26.63531	15.07317	-1.767067	0.0785
GCPI(-1)	27.68259	24.43360	1.132972	0.2584

R-squared 0.083821 Mean dependent var 1.197711
Adjusted R-squared 0.072125 S.D. dependent var 1.467440
S.E. of regression 1.413530 Akaike info criterion 3.546652
Sum squared resid 469.5459 Schwarz criterion 3.604836
Log likelihood -419.8250 Hannan-Quinn criter 3.570099
F-statistic 7.166732 Durbin-Watson stat 0.036049
Prob(F-statistic) 0.000126

表3-14 广义差分模型的最小二乘回归结果

Dependent Variable: DR
Method: Least Squares
Date: 02/18/22 Time: 19:50
Sample (adjusted): 2002M02 2021M12
Included observations: 239 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.004632	0.023972	-0.193217	0.8470
DIP	0.010792	0.011403	0.946486	0.3449
DGM2	-5.305996	2.584050	-2.053365	0.0411
DGCPI(-1)	6.636457	3.317794	2.000262	0.0466

R-squared 0.061864 Mean dependent var 0.014505
Adjusted R-squared 0.049888 S.D. dependent var 0.190236
S.E. of regression 0.185430 Akaike info criterion -0.515685
Sum squared resid 8.080299 Schwarz criterion -0.457501
Log likelihood 65.62433 Hannan-Quinn criter -0.492238
F-statistic 5.165599 Durbin-Watson stat 1.792711
Prob(F-statistic) 0.001784

例3.2 利率的变化

2. 自相关的修正——杜宾两步法

第一步，对模型

$$R_t - \rho R_{t-1} = (1 - \rho)b_0 + b_1(IP_t - \rho IP_{t-1}) + b_2(GM2_t - \rho GM2_{t-1}) + b_3(GCPI_{t-1} - \rho GCPI_{t-2}) + \varepsilon_t - \rho \varepsilon_{t-1}$$

进行变换，得到：

$$R_t = (1 - \rho)b_0 + \rho R_{t-1} + b_1 IP_t - \rho b_1 IP_{t-1} + b_2 GM2_t - \rho b_2 GM2_{t-1} + b_3 GCPI_{t-1} - \rho b_3 GCPI_{t-2} + v_t$$

对上式用OLS估计：

Equation Estimation

Specification Options

Equation specification

Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)*X$.

r c r(-1) ip ip(-1) gm2 gm2(-1) gcpi(-1) gcpi(-2)

Estimation settings

Method: LS - Least Squares (NLS and ARMA)

Sample: 2002M01 2021M12

确定 取消

例3.2 利率的变化

2. 自相关的修正——杜宾两步法

R_t 的估计系数就是自相关系数 ρ 的估计值: $\hat{\rho} = 0.991885$

表3-15 杜宾两步法估计自相关系数 ρ 结果

Dependent Variable: R					
Method: Least Squares					
Date: 02/18/22 Time: 19:53					
Sample (adjusted): 2002M03 2021M12					
Included observations: 238 after adjustments					
Variable	Coefficient	Std. Error	t-Statistic	Prob.	
C	0.165518	0.262343	0.630923	0.5287	
R(-1)	0.991885	0.008599	115.3462	0.0000	
IP	0.002952	0.012132	0.243318	0.8080	
IP(-1)	-0.004366	0.012062	-0.361955	0.7177	
GM2	-8.048690	3.000452	-2.682493	0.0078	
GM2(-1)	3.752337	2.704166	1.387614	0.1666	
GCPI(-1)	6.285715	3.775415	1.664907	0.0973	
GCPI(-2)	-6.660381	3.773208	-1.765177	0.0789	
R-squared	0.984576	Mean dependent var		1.195366	
Adjusted R-squared	0.984107	S.D. dependent var		1.470084	
S.E. of regression	0.185332	Akaike info criterion		-0.500302	
Sum squared resid	7.900013	Schwarz criterion		-0.383587	
Log likelihood	67.53597	Hannan-Quinn criter.		-0.453264	
F-statistic	2097.412	Durbin-Watson stat		1.862402	
Prob(F-statistic)	0.000000				

例3.2 利率的变化

2. 自相关的修正——杜宾两步法

第二步，用 $\hat{\rho} = 0.991885$ 对原始数据进行差分变换，余下步骤与已知 ρ 的情况下采用广义差分法修正自相关相同。最终得到的结果如表3-16所示，与广义差分法的结论类似。

表3-16 杜宾两步法修正自相关后的回归结果

Dependent Variable: DR_2

Method: Least Squares

Date: 02/18/22 Time: 20:00

Sample (adjusted): 2002M02 2021M12

Included observations: 239 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.006108	0.015355	-0.397795	0.6911
DIP_2	0.010499	0.011322	0.927278	0.3547
DGM2_2	-5.321121	2.562973	-2.076152	0.0390
DGCPI_2(-1)	6.646814	3.289592	2.020559	0.0445

R-squared	0.062525	Mean dependent var	0.002565
Adjusted R-squared	0.050558	S.D. dependent var	0.189603
S.E. of regression	0.184748	Akaike info criterion	-0.523055
Sum squared resid	8.020964	Schwarz criterion	-0.464872
Log likelihood	66.50508	Hannan-Quinn criter.	-0.499609
F-statistic	5.224483	Durbin-Watson stat	1.822800
Prob(F-statistic)	0.001650		



本章小结

- 在金融计量和经济计量诸多分析中都要面对异方差问题，异方差问题是金融计量和经济计量时不满足经典回归条件的几个主要问题之一。本章首先明确了异方差的定义，并简要说明了其产生原因及后果，在此基础上从图示法和解析法两个方面介绍了诸多异方差的检验方法，然后具体介绍了修正异方差的方法，并辅以实例详细说明了异方差检验到修正的过程。
- 另外，作为经典线性回归模型（**CLRM**）五个假设的有一个破坏——自相关，本章梳理了自相关问题的解决思路。其中，观测是否存在自相关，可以选择图示法或者解析法；如何解决自相关问题，可以通过广义差分法或者杜宾两步法等等。



本章小结

- 异方差的定义，产生原因及后果
- 异方差的检验方法
- 异方差的修正方法
- 自相关的产生原因
- 忽略自相关的严重后果
- 自相关的检验
- 自相关的修正