第六章 自相关

引子　农村居民的边际消费倾向是什么?

研究农村居民人均消费Y与人均纯收入X的关系，设定模型：

$$Y_t = \beta_1 + \beta_2 X_t + u_t$$

用普通最小二乘法估计其参数，结果为

$$\hat{Y}_t = -11.1195 + 0.7835 X_t$$

（23.4635）(0.0117)

t= (-0.4739) (66.9993)

$R^2 = 0.9947$　F=4488.91

检验结果：回归系数标准误差非常小，t统计量很大，说明农村居民人均纯收入X对人均消费Y的影响非常显著。同时可决系数也非常高，F统计量=4122.531，也表明模型异常的显著。但标准误差之低以及t统计量和F统计量之高，让人难以置信。那么，有什么理由提出这样的质疑呢?怎样估计更为接近实际的边际消费倾向呢？

本章讨论四个问题：

- 自相关的概念和产生的原因
- 自相关的后果
- 自相关的检验方法
- 自相关的补救方法

第一节 自相关的概念

一、什么是自相关

一般概念:自相关是指以时间和空间为顺序的观测值序列中各部分之间的相关关系，也称序列相关。

计量经济学中的概念：特指随机扰动项逐次观测值相互之间的相关关系。

一般表示为: $$Cov(u_i, u_j) = E(u_i u_j) \neq 0 \quad (i \neq j)$$

自相关程度的度量

自相关系数

$$\rho = \frac{\sum_{t=2}^{n} u_t u_{t-1}}{\sqrt{\sum_{t=2}^{n} u_t^2 \sum_{t=2}^{n} u_{t-1}^2}}$$

自相关的形式

如果$Cov(u_t, u_{t-1}) \neq 0$ 称 u_t 序列存在一阶自相关

如果u_t的自相关形式为: $u_t = \rho u_{t-1} + \varepsilon_t$

其中: ε_t 满足OLS基本假定:

$$E(\varepsilon_t) = 0 \quad Var(\varepsilon_t) = \sigma_\varepsilon^2 \quad Cov(\varepsilon_t, \varepsilon_s) = 0 \quad (t \neq s)$$

称 u_t 呈现一阶自回归形式

称为一阶自回归系数，近似于一阶自相关系数 $|\rho| \leq 1$

因为 $$\hat{\rho} = \frac{\sum_{t=2}^{n} u_t u_{t-1}}{\sum_{t=2}^{n} u_{t-1}^2} \approx \frac{\sum_{t=2}^{n} u_t u_{t-1}}{\sqrt{\sum_{t=2}^{n} u_t^2 \sum_{t=2}^{n} u_{t-1}^2}}$$

（回归系数公式）（相关系数公式）

在样本容量大时有
$$\sum u_t^2 \approx \sum u_{t-1}^2$$

也可能是二阶自回归形式，可记为 $AR(2)$

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \varepsilon_t$$

u_t 的K阶自回归形式，可记为 $AR(k)$

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_k u_{t-k} + \varepsilon_t$$

自回归的形式将在时间序列中讨论。

这里只讨论一阶自回归形式的自相关问题

- 一阶自回归形式较为简单
- 在实际计量分析中处理一阶自回归形式常能取得较好效果.

一阶自回归形式的自相关性质

对于 $u_t = \rho u_{t-1} + \varepsilon_t$ 可以证明：

$$
\begin{aligned}
u_t &= \rho(\rho u_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= \rho^2(\rho u_{t-3} + \varepsilon_{t-2}) + \rho\varepsilon_{t-1} + \varepsilon_t \\
&= \rho^3(\rho u_{t-4} + \varepsilon_{t-3}) + \rho^2\varepsilon_{t-2} + \rho\varepsilon_{t-1} + \varepsilon_t \\
&= \cdots
\end{aligned}
$$

一般关系： $$u_t = \rho u_{t-1} + \varepsilon_t = \sum_{k=0}^{\infty} \rho^k \varepsilon_{t-k}$$

期望 $$E(u_t) = \sum_{k=0}^{\infty} \rho^k E(\varepsilon_{t-k}) = 0$$

方差：

$$\sigma_u^2 = Var(u_t) = Var(\sum_{t=0}^{\infty} \rho^k \varepsilon_{t-k}) = \sum_{t=0}^{\infty} \rho^{2k} Var(\varepsilon_{t-k})$$

$$= \sigma_\varepsilon^2 (1 + \rho^2 + \rho^4 + \cdots) = \frac{\sigma_\varepsilon^2}{1-\rho^2}$$

协方差：

$k = 1$时

$$Cov(u_t, u_{t-1}) = \rho(\sigma_\varepsilon^2 + \rho^2 \sigma_\varepsilon^2 + \rho^4 \sigma_\varepsilon^2 + \cdots) = \rho \frac{\sigma_\varepsilon^2}{1-\rho^2} = \rho \sigma_u^2$$

类推可得

$k = 2$时

$$Cov(u_t, u_{t-2}) = E(u_t u_{t-2}) = \rho^2 \frac{\sigma_\varepsilon^2}{1-\rho^2} = \rho^2 \sigma_u^2$$

$$Cov(u_t, u_{t-k}) = E(u_t u_{t-k}) = \rho^k \frac{\sigma_\varepsilon^2}{1-\rho^2} = \rho^k \sigma_u^2$$

自相关产生的原因

（1）经济变量本身的惯性作用

经济变量与前几个时期的数值往往有关，如本期消费常与前期消费有关

（2）经济行为本身的滞后性

如本期消费还依赖于前期收入，而前期收入未纳入模型

（3）设定偏倚（虚假自相关）

如省略重要解释变量、不正确的函数形式可引起自相关

（4）数据的加工引起自相关

如数据修匀平滑， 用内插和外推取得数据

（5）一些其他的经济问题

某些偶然因素如灾害、政治因素的长期影响、蛛网现象等

第二节 自相关的后果

一、对参数估计的影响

1. 参数的OLS估计式仍然是无偏的 (无偏性证明中未涉及自相关)

2. 用OLS估计的参数的方差不再具有最小方差 (可以找到比OLS更小方差的估计式)

存在自相关时仍用经典假定下公式可能严重低估真实方差

$$Var(\hat{\beta}_2) < Var(\hat{\beta}_2^*)$$

其中 $Var(\hat{\beta}_2)$ 是经典假定下公式计算的方差

$Var(\hat{\beta}_2^*)$ 是存在自相关时所估计参数的真实方差

3. 用 $\sum e_i^2$ 估计 u_i 的方差，会低估 u_i 的真实方差

（可以证明） $\hat{\sigma}^2 = \Sigma e_i^2 / (n-k)$ 将低估真实的 σ^2

回顾：异方差和自相关对方差的影响

对于 $Y_i = \beta_1 + \beta_2 X_i + u_i$

由
$$Var(\hat{\beta}_2) = \frac{\sum x_i^2 E(u_i^2) + 2\sum_{i \neq j} x_i x_j E(u_i u_j)}{(\sum x_i^2)^2}$$

在同方差且无自相关时

$[E(u_i^2) = \sigma^2, E(u_i u_j) = 0]$

$$Var(\hat{\beta}_2) = \frac{\sum x_i^2 E(u_i^2)}{(\sum x_i^2)^2} = \frac{\sigma^2}{\sum x_i^2}$$

在异方差但无自相关时

$[E(u_i^2) = \sigma_i^2, E(u_i u_j) = 0]$

$$Var(\hat{\beta}_2) = \frac{\sum x_i^2 E(u_i^2)}{(\sum x_i^2)^2} = \frac{\sum x_i^2 \sigma_i^2}{(\sum x_i^2)^2}$$

在同方差但自相关时

$[E(u_i^2) = \sigma^2, E(u_i u_j) \neq 0]$

$$Var(\hat{\beta}_2) = \frac{\sigma^2}{\sum x_i^2} + \frac{2\sum_{i \neq j} x_i x_j E(u_i u_j)}{(\sum x_i^2)^2}$$

由于$E(u_i u_j)=?$未知，$Var(\hat{\beta}_2)$的估计出现困难

因为 $Cov(u_t, u_{t-k}) = E(u_t u_{t-k}) = \rho^k \dfrac{\sigma_\varepsilon^2}{1-\rho^2} = \rho^k \sigma_u^2$

$$Var(\hat{\beta}_2) = \frac{\sigma_u^2}{\sum x_i^2} + \frac{2\sum_{i\neq j} x_i x_j E(u_i u_j)}{(\sum x_i^2)^2} = \frac{\sigma_u^2}{\sum x_i^2} + \frac{2\sum_{t=1}^{n-1}\sum_{k=1}^{n-1} x_t x_{t-k} E(u_t u_{t-k})}{(\sum x_i^2)^2}$$

$$= \frac{\sigma_u^2}{\sum x_i^2} + \frac{2\sum_{t=1}^{n-1}\sum_{k=1}^{n-1} x_t x_{t-k} \rho^k \sigma_u^2}{(\sum x_t^2)^2} = \frac{\sigma_u^2}{\sum x_t^2} + \left[\frac{2\sigma_u^2}{\sum x_t^2} \frac{\sum_{t=1}^{n-1}\sum_{k=1}^{n-1} \rho^k x_t x_{t-k}}{\sum x_t^2}\right]$$

$$= \frac{\sigma_u^2}{\sum x_t^2}\left\{1 + 2\rho \frac{\sum_{t=1}^{n-1} x_t x_{t+1}}{\sum x_t^2} + 2\rho^2 \frac{\sum_{t=1}^{n-2} x_t x_{t+2}}{\sum x_t^2} + \cdots + 2\rho^{n-1} \frac{x_1 x_n}{\sum x_t^2}\right\}$$

真实方差： $$Var(\hat{\beta}_2)=\frac{\sigma_u^2}{\sum x_t^2}+[\frac{2\sigma_u^2}{\sum x_t^2}\frac{\sum_{t=1}^{n-1}\sum_{k=1}^{n-1}\rho^k x_t x_{t-k}}{\sum x_t^2}]$$

- 存在自相关时 $\rho \neq 0$，在经济问题中常见的是 $\rho > 0$，

且解释变量经常正自相关，交叉项 $x_t x_{t-k}$ 为正，大多数经济应用中 $\sum_{t=1}^{n-1}\sum_{k=1}^{n-1}\rho^k x_t x_{t+k}$ 为正。通常只用 $Var(\hat{\beta}_2)=\frac{\sigma_u^2}{\sum x_i^2}$ 会低估OLS估计量的真实方差。

- 如果 $\rho < 0$，k 为奇数时 $\rho^k < 0$，k 为偶数时 $\rho^k > 0$，

$\sum_{t=1}^{n-1}\sum_{k=1}^{n-1}\rho^k x_t x_{t+k}$ 的符号难以断定，用 $\frac{\sigma_u^2}{\sum x_i^2}$ 也可能高估OLS估计量的真实方差，但对OLS估计量方差的估计也是有偏的。

真实方差： $Var(\hat{\beta}_2) = \frac{\sigma_u^2}{\sum x_t^2} + [\frac{2\sigma_u^2}{\sum x_t^2} \frac{\sum_{t=1}^{n-1}\sum_{k=1}^{n-1} \rho^k x_t x_{t+k}}{\sum x_t^2}]$

用 $\sum e_i^2$ 还会低估 u_t 的真实方差，因为证明见教材p159(附录)

$$E(\sum e_i^2) = \sigma^2[(n-2) - (2\rho \frac{\sum X_t X_{t+1}}{\sum X_t^2} + 2\rho^2 \frac{\sum X_t X_{t+2}}{\sum X_t^2} + \cdots + 2\rho^{n-1} \frac{\sum X_t X_n}{\sum X_t^2})]$$

经济问题中自相关时通常为正值

只用 $\sum e_i^2/(n-2)$ 会过低估计 σ_u^2 。

这样，将会进一步低估 $\hat{\beta}_2$ 的真实方差，因为在低估 σ_u^2 的基础上用 $Var(\hat{\beta}_2) = \sigma_u^2/\sum x_i^2$ 可能更加过低估计参数真实方差。

注意关系：存在自相关时OLS非有效是指真实方差不再具有最小方差性，而真实方差可能会被标准方差公式低估。

对模型检验的影响

1. 参数的显著性检验将失效

可能过低估计参数真实方差和标准误差 $SE(\hat{\beta}_2)$

则可能过高估计$t=\hat{\beta}_2 / SE(\hat{\beta}_2)$，而夸大$\beta_2$的显著性，使得 t 检验失效，同理，F 检验也将失效

2. 区间估计变得无意义

由于方差标准误差的估计是有偏的，或被过低估计，

区间估计不可信，变得无意义。

3. 对模型预测的影响

模型预测的精度决定于：◆抽样误差◆u_i的方差σ^2

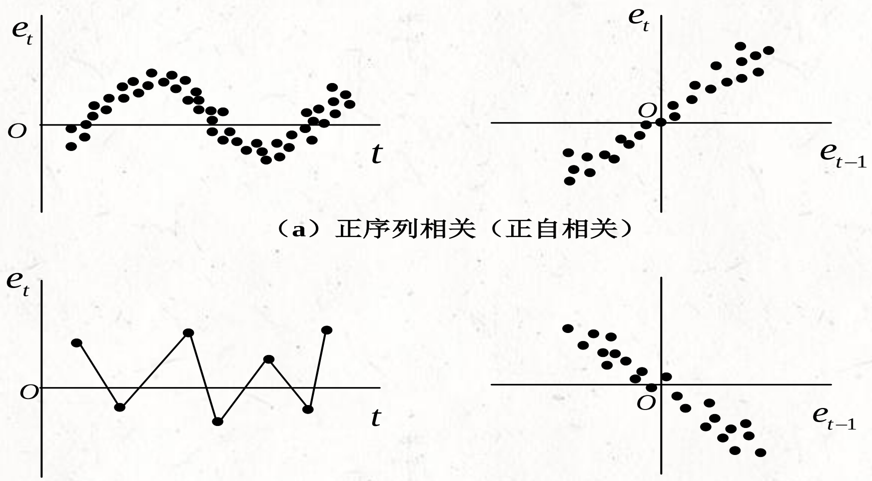
◆抽样误差来自于对$\hat{\beta}_j$的估计，存在自相关时，OLS估计的 $Var(\hat{\beta}_j)$ 不再最小，会影响抽样误差。

◆在自相关情形下，用$\hat{\sigma}^2 = \Sigma e_i^2 / n-k$ 对 σ^2的估计也会不可靠。

影响预测精度的两个因素都可能因自相关的存在而加大不确定性，使得预测的置信区间不可靠，预测精度下降。

一、图解法

用样本回归剩余e_i代替u_i，绘制以e_i为纵坐标，以 e_{i-1} 或时间顺序 t 为横坐标的坐标图，观测是否存在自相关，如



（b）负序列相关（负自相关）

二、德宾—沃森检验（Durbin—Watson检验）

1. 基本思想：

将e_i 视为对u_i的估计，寻求适当的检验统计量

原假设：$H_0: \rho = 0 \quad H_1: \rho \neq 0$

建立 DW 统计量（也称d统计量)：

$$d = \frac{\sum_{2}^{n}(e_t - e_{t-1})^2}{\sum_{1}^{n} e_t^2}$$

关键是设法确定D的分布 。

可以证明: $d=\dfrac{\sum_{2}^{n}(e_t-e_{t-1})^2}{\sum_{1}^{n}e_t^2}=\dfrac{\sum e_t^2+\sum e_{t-1}^2-2\sum e_t e_{t-1}}{\sum e_t^2}$

大样本时: $\sum e_{t-1}^2 \approx \sum e_t^2$ （只差一次观测的 e_i^2 ）

$$d \approx 2\left(1-\frac{\sum e_t e_{t-1}}{\sum e_t^2}\right) \approx 2(1-\hat{\rho})$$

$$\hat{\rho}=\frac{\sum u_t u_{t-1}}{\sum u_t^2} \approx \frac{\sum e_t e_{t-1}}{\sum e_t^2}$$

可见，对ρ=0的检验等价于对 d=2 的检验

2. 德宾—沃森DW检验的假定条件：

（1）解释变量非随机

（2）模型包括截距项（不是通过原点的回归）

（3）解释变量中不含滞后被解释变量，如 Y_{t-1}

（4）u_i 的自相关是一阶自回归形式，即

$$u_t = \rho u_{t-1} + \varepsilon_t$$

（5）无缺损数据

3. 具体做法

(1)　进行OLS回归得剩余 e_i

(2)　计算统计量

$$d=\frac{\sum_2^n(e_t-e_{t-1})^2}{\sum_1^n e_t^2}$$

(3) 确定d 的概率分布：它与 X_i、样本容量 n 、解释变量个数 k’ 都有关，具体确定其分布性质很困难。

但D-W给出了d统计量有价值的临界值（d统计量表）

(4) 给定显著性水平α ，查D—W 的d统计量表，得与样本容量为 n ，解释变量个数为 k’ 对应的临界值 d_L 和 d_U

(5) 判断是否存在自相关

临界值 d_L 和 d_U 把d值分为五个区域:(见下页)

假设：H_0：无正自相关 或 H_0^*：无负自相关

d_L 和 d_U 把d 值可分为五个区域：

正自相关区域 拒绝 H_0	无结论区域	无自相关区域 不拒绝 H_0 和 H_0^*	无结论区域	负自相关区域 拒绝 H_0^*

$0 \qquad d_L \qquad d_U \qquad 2 \qquad 4-d_U \qquad 4-d_L \qquad 4$

判断：

（1）$0 < d < d_L$ 时，拒绝 H_0，存在正自相关

（2） $d_L < d < d_U$ 时，不能确定是否存在自相关

（3）$d_U < d < 4-d_U$ 时，不拒绝 H_0 和 H_0^*，不存在自相关

（4） $4-d_U < d < 4-d_L$ 时，不能确定是否存在自相关

（5） $4-d_L < d < 4$ 时，拒绝 H_0^*，存在负自相关

4. DW检验的优点和局限

优点：依据通常要计算的 e_i ，使用方便

局限:（1）有假定前提条件（ 5个条件）

（2）要求有足够样本量（一般要求n≥15）

(3）有不确定区域

修订方式:

- $d_U < d < 4 - d_U$ 时，接受 $H_0: \rho = o$ ，认为不存在自相关
- $d < d_U$ 或 $d > 4 - d_U$ 就拒绝 $H_0: \rho = o$ ，认为存在自相关

（这是扩大拒绝区域，不确定时宁可拒绝而不宜接受的"宁左勿右"的作法)

三、(Breusch-Godfrey)LM检验法

LM (BG)检验是Breusch和Godfre基于拉格朗日乘数（LM）原理，提出的检验自相关的方法（BG）

$$Y_t = \beta_1 + \beta_2 X_{2t} + \cdots + \beta_k X_{kt} + u_t$$

LM检验不只限于检验一阶自相关，即

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_p u_{t-p} + v_t$$

其中 v_t 满足古典假定。

原假设：　$H_0: \rho_1 = \cdots = \rho_p = 0$

具体步骤：

I. 用OLS估计原模型，得到残差 e_t

II. 用残差 e_t 对解释变量X及滞后残差 e_{t-j} 作辅助回归，即

$$e_t = \alpha_1 + \alpha_2 X_{2t} + \cdots + \alpha_k X_{kt} + \rho_1 e_{t\text{-}1} + \rho_2 e_{t\text{-}2} + \cdots + \rho_p e_{t\text{-}p} + v_t$$

III. 计算检验统计量 $\mathrm{LM}=T\cdot R^2$，其中T为辅助回归实际数据个数，R^2为辅助回归可决系数。大样本下：

$$LM \sim \chi_p^2$$

为了避免残差取滞后而丧失有效样本，不影响LM统计量的渐近性，将n个样本之前的p期初始值都预处理为0（此时，有效样本T=n）

Ⅳ. 检验

给定显著性水平α，查χ^2分布表得临界值$\chi_\alpha^2(p)$，

- 如果$LM \geq \chi_\alpha^2(p)$，H_0不合理，则拒绝原假设H_0，即认为至少有一个ρ不为0，即存在自相关
- 如果$LM < \chi_\alpha^2(p)$，则不拒绝H_0，即认为模型中随机误差无自相关。

LM(BG)检验的特点

- LM(BG)检验不止限于一阶自相关，还适合于高阶自相关
- 适合模型中的解释变量含有滞后被解释变量的情况
- LM(BG)检验的滞后长度p不能先验确定。

实际中，可逐次向高阶检验，并结合辅助回归中滞后项参数的显著性帮助判断自相关的阶数。或者应该设置一个合理的最大阶数上限，然后选择试AIC/SIC值达到最小的阶数。

第四节 自相关的补救办法

一、纠正设定误差　　　可减弱自相关

设定误差造成的自相关，只能通过改变模型的设定去消除

1.引入导致自相关的省略解释变量

1）发现和确认引起自相关的解释变量（如滞后变量）

可将剩余 e_i 对省略的主要解释变量逐个回归

2）将确认的变量引入模型，消除或减轻自相关

2. 改变导致自相关的函数形式

1）发现错误的函数形式

用剩余 e_i 对解释变量较高次幂回归，检验新剩余是否还有自相关

2）改变函数形式，减弱自相关影响

注意：如果是真实自相关，纠正设定误差方法无效

二、已知自相关系数ρ时对模型的变换

当u_t为一阶自相关形式，并已知ρ时，可用广义差分法

基本思想： 原模型 $Y_t = \beta_1 + \beta_2 X_t + u_t$

因为 $u_t = \rho u_{t-1} + \varepsilon_t$ ，已知 $\varepsilon_t = u_t - \rho u_{t-1}$ 无自相关，

可设法将模型的扰动项变换为 ε_t ，即广义差分形式

方法：用“（原模型）— $\rho\times$（滞后一个期的模型）”得

$$\underline{Y_t - \rho Y_{t-1}} = \underline{(\beta_1 - \rho\beta_1)} + \beta_2 \underline{(X_t - \rho X_{t-1})} + \underline{(u_t - \rho u_{t-1})}$$

$$Y_t^* = \beta_1^* + \beta_2 X_t^* + \varepsilon_t$$

ε_t 满足基本假定： 零均值 $E(\varepsilon_t) = 0$

同方差 $Var(\varepsilon_t) = \sigma_\varepsilon^2$

无自相关 $\mathrm{cov}(\varepsilon_t, \varepsilon_s) = 0 \quad (t \neq s)$

具体方法

估计变换后的模型，得β_1^* 和β_2，再由ρ可计算出 β_1:

因为 $\beta_1^* = \beta_1 - \rho\beta_1$ 则 $\beta_1 = \beta_1^* / (1-\rho)$

注意:

- 前提条件是已知自相关系数 ρ
- 广义差分后只有n-1个观测值,为避免观测值损失，*Y*和*X*的第一个观测值可用如下普莱斯—温斯腾变换得第一个观测值

$$Y_1^* = Y_1\sqrt{1-\rho^2} \qquad X_1^* = X_1\sqrt{1-\rho^2}$$

（其他解释变量用同样方法变换得第一个观测值）

- 模型已成为变换了的新变量之间的回归

三、自相关系数ρ未知时模型的变换

思想：通常ρ未知，为用模型变换处理自相关，必须 设法找到ρ的估计值

方法1. 用dw 统计量估计ρ

在大样本时

已知 $d \approx 2(1-\hat{\rho})$

因此 $\hat{\rho} \approx 1 - d/2$

从DW检验中已得到d 统计量，即可估计出 $\hat{\rho}$

注意：此方法只有在大样本时才有效

方法2：用残差 e_i 直接估计ρ

思想:由于一阶自回归系数，近似于一阶自相关系数

$$\hat{\rho}=\sum_{t=2}^{n}u_t u_{t-1}\Big/\sum_{t=2}^{n}u_{t-1}^2\approx\sum_{t=2}^{n}u_t u_{t-1}\Big/\sqrt{\sum_{t=2}^{n}u_t^2\sum_{t=2}^{n}u_{t-1}^2}$$

用e_i替代u_i去估计ρ

原模型作OLS估计，计算 e_i

作过原点的回归 $e_t=\beta e_{t-1}+u_t$

在Eviews中生成新变量e=resid，在命令栏输入“ls e e（-1）”/回车，得到估计的$\hat{\beta}$

可视为估计的 $\hat{\beta}\approx\hat{\rho}\approx(\sum e_t e_{t-1})/(\sum e_{t-1}{}^2)$

方法3. 科克兰（Cochrane）—奥卡特（Orcutt）迭代法

基本思想：利用剩余 e_i 去获得未知的 ρ 的信息。

通过逐次迭代寻求（逐步逼近）更满意的 ρ 的估计值

▼ 原模型 $Y_t = \beta_t + \beta_2 X_t + u_t$ 且 $u_t = \rho u_{t-1} + v_t$

可用剩余 e 替代 u 去估计 ρ

方法:作回归 $e_t = \rho e_{t-1} + v_t$ $\rho^* = \sum e_t e_{t-1} / \sum e_{t-1}^2$

▼用估计的 ρ^* 对原模型作广义差分回归，得剩余项 $e_{(1)}^*$

▼由所得剩余 $e_{(1)}^*$ 重新估计 ρ_1^*，再用 ρ_1^* 对原模型作广义差分回归，得剩余项 $e_{(2)}^*$

▼用剩余 $e_{(2)}^*$ 再估计 ρ_2^*，又用 ρ_2^* 对原模型作广义差分回归直到估计的 ρ 收敛满足精度要求，或回归所得DW统计量通过零假设（不存在自相关）为止。

迭代的方法步骤：

1）用OLS估计原模型，计算回归剩余e_t，并估计ρ^*

$$\rho^* = (\sum e_t e_{t-1}) / (\sum e_{t-1}^2)$$

2）用 ρ^* 作一阶差分回归

$$Y_t - \rho^* Y_{t-1} = \beta_1(1-\rho^*) + \beta_2(X_t - \rho^* X_{t-1}) + \varepsilon_t^{(1)}$$

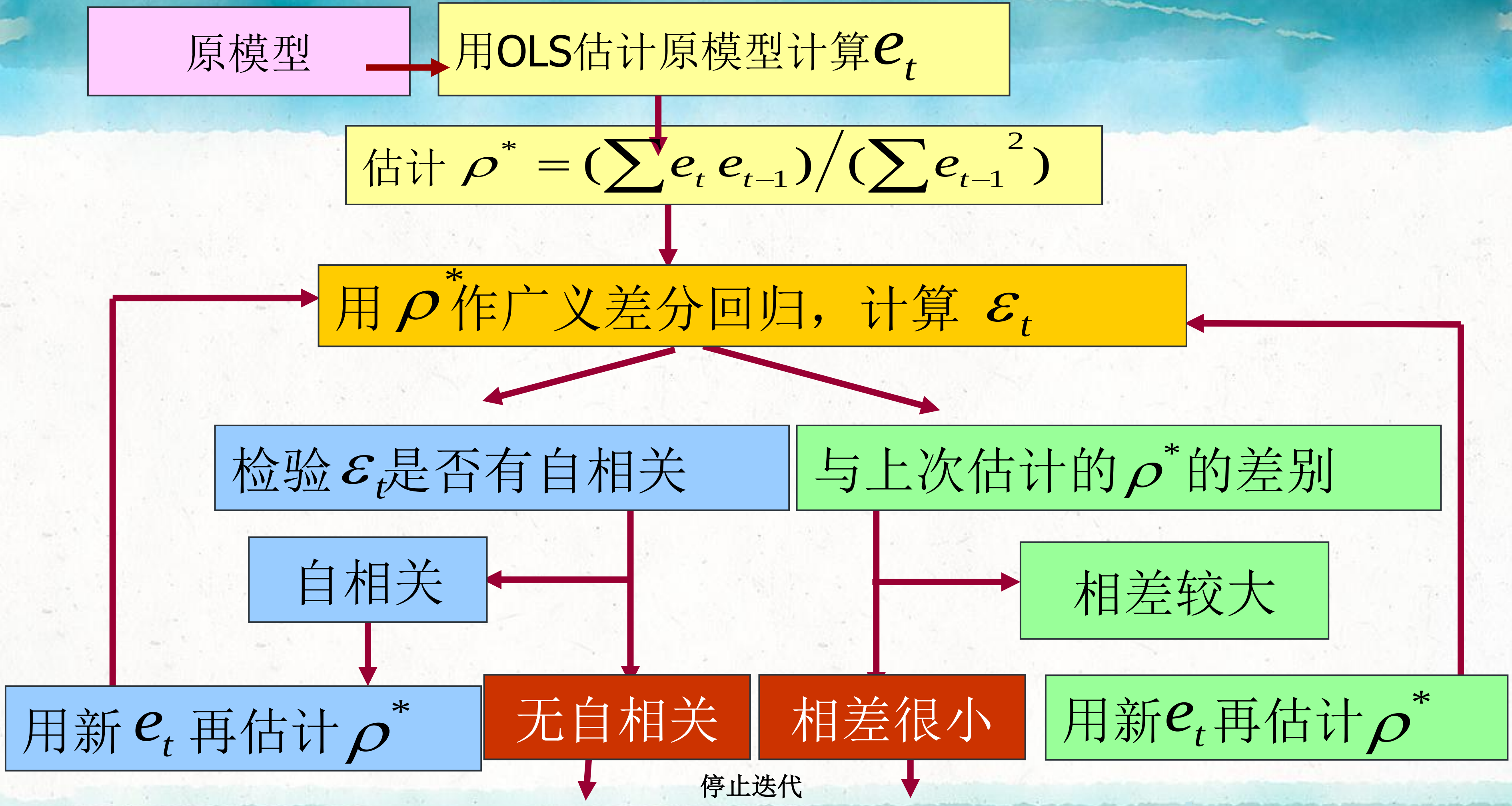
检验 $\varepsilon_t^{(1)}$ 的自相关性，若还有自相关，用 $e_t^{(1)}$ 第二次估计 ρ_1^*

$$\rho_1^* = (\sum e_t^{(1)} e_{t-1}^{(1)}) / (\sum e_{t-1}^{(1)2})$$

3）用 ρ_1^* 作一阶差分回归

$$Y_t - \rho_1^* Y_{t-1} = \beta_1(1-\rho_1^*) + \beta_2(X_t - \rho_1^* X_{t-1}) + \varepsilon_t^{(2)}$$

检验 $\varepsilon_t^{(2)}$ 的自相关性，若无自相关，迭代停止，得到 β_1, β_2 的估计值。若还有自相关，再用 $e_t^{(2)}$ 第三次估计 ρ_2^* ，继续广义差分回归，直到经检验无自相关为止。


原模型
用OLS估计原模型计算e_t
估计 $\rho^* = (\sum e_t e_{t-1}) / (\sum e_{t-1}^2)$
用 ρ^*作广义差分回归，计算 ε_t
检验ε_t是否有自相关
与上次估计的ρ^*的差别
自相关
相差较大
用新e_t再估计ρ^*
无自相关
相差很小
用新e_t再估计ρ^*
停止迭代

方法4. 德宾两步法

原模型 $Y_t = \beta_t + \beta_2 X_t + u_t$

基本思想和作法：

设法间接地估计出ρ，再利用$\hat{\rho}$ 作广义差分变换原模型

1）如果已知ρ ，可对原模型作广义差分变换

$$Y_t - \underline{\rho Y_{t-1}} = (\beta_1 - \rho\beta_1) + \beta_2(X_t - \rho X_{t-1}) + (u_t - \rho u_{t-1})$$

2）将上式中 ρY_{t-1} 移项到方程右边

$$Y_t = \beta_1(1-\rho) + \underline{\rho Y_{t-1}} + \beta_2 X_t - \beta_2\rho X_{t-1} + (u_t - \rho u_{t-1})$$

其中$u_t - \rho u_{t-1}$ 满足基本假定，无自相关

3) 可用OLS法估计上式，估计出$\hat{\rho}$，它是ρ 的一致估计式。

（以上为第一步）

4) 用估计的 $\hat{\rho}$ 对原模型作广义差分变换，并用OLS估计其参数，得原模型参数估计值。（第二步）

案例：中国农村居民收入－消费模型

研究范围：中国农村居民收入－消费（1990~2015）

研究目的：改革开放以来，中国农村居民的收入和消费支出都在快速增长。2015年农村居民人均可支配收入11422元，农村居民人均消费9223元，中国乡村人口占总人口的43.90%，农村居民收入和消费的状况是令人关注的经济问题。

建立模型：

$$Y_t = \beta_1 + \beta_2 X_t + u_t$$

Yt－农村居民人均消费，Xt－农村居民人均收入。

（数据见P148表6.3）

使用普通最小二乘法估计消费模型得

Dependent Variable: Y
Method: Least Squares
Date: 10/30/19 Time: 11:14
Sample: 1990 2015
Included observations: 26

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-11.11948	23.46353	-0.473905	0.6399
X	0.783541	0.011695	66.99933	0.0000

R-squared	0.994682	Mean dependent var	1346.764
Adjusted R-squared	0.994460	S.D. dependent var	809.9703
S.E. of regression	60.28520	Akaike info criterion	11.10985
Sum squared resid	87223.33	Schwarz criterion	11.20663
Log likelihood	-142.4281	Hannan-Quinn criter.	11.13772
F-statistic	4489.911	Durbin-Watson stat	0.508796
Prob(F-statistic)	0.000000		

t 和 F很显著

但DW表明

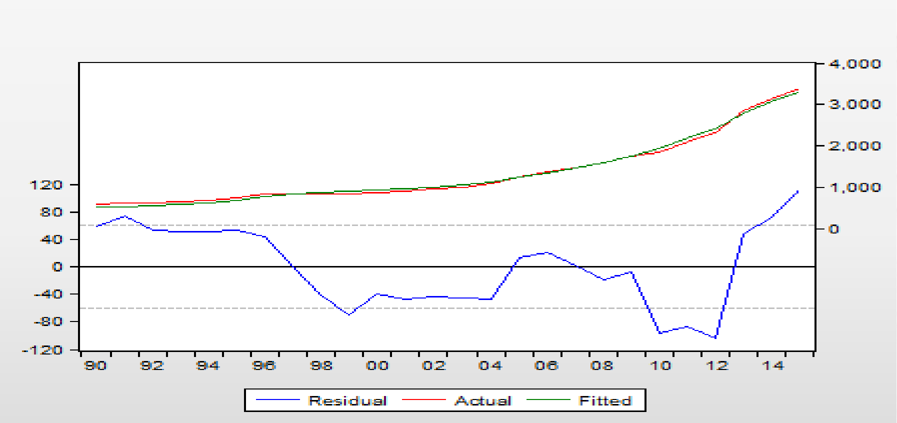
可能有自相关

注：为消除价格变动因素的影响，这里使用经消费价格指数进行调整后的1990年可比价格计算的人均纯收入和人均消费支出数据进行回归。

自相关检验

该回归方程F统计量为4488.91， 回归系数的t检验很显著。对样本量为26、一个解释变量的模型、5%显著水平， 查*DW*统计表可知， dL=1.302， dU= 1.461， 模型中*DW*=0.5088<*dL*， 显然消费模型中有正自相关。

点击EViews方程输出窗口的按钮Resids可得到残差图，从残差 e_t 与时间 t 的图中也看出存在正自相关，

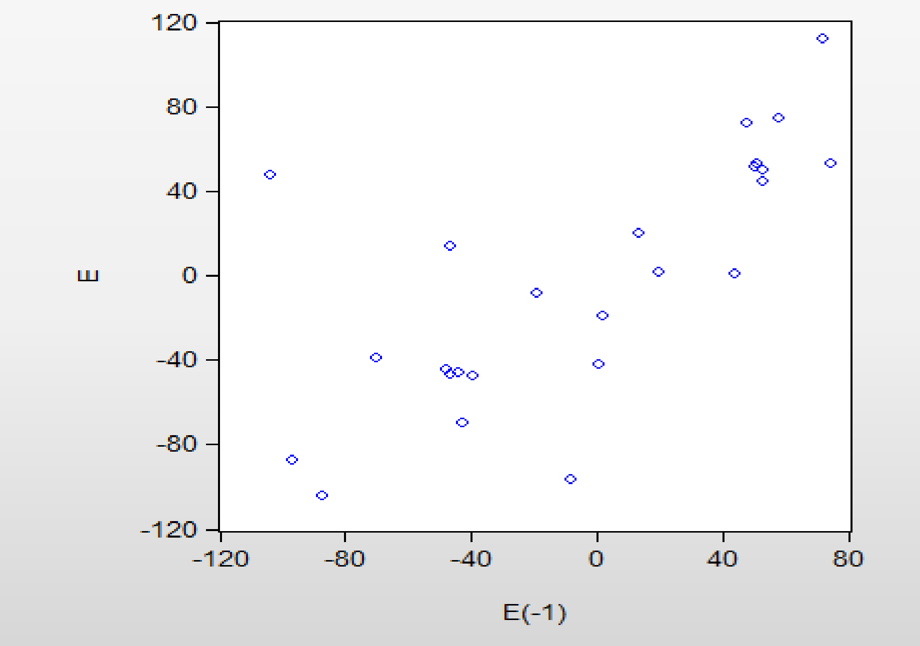


或者作残差 e_t 与 e_{t-1} 的图形，从图中也看出可能存在正自相关。

在命令窗中依次输入：

"genr e=resid"
"scat e(-1) e"

注：最好不要直接使用工作文件中默认的序列resid。



BG（LM）检验

在回归输出结果中点击“View/Residual Diagnostics /Series Correlation LM Test”，在“lags to includes”中选取滞后阶数，如“2”，回车即得BG检验结果。

LM=26×0.545376=14.17977，其*p*值为0.0008，也表明存在自相关。

注：可以设置一个合理的最大阶数上限，然后选择使AIC和SIC值达到最小的阶数（即，每次在“lags to includes”中填入不同滞后阶数尝试，看哪个阶数的AIC和SIC值最小。

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	13.19581	Prob. F(2,22)	0.0002
Obs*R-squared	14.17977	Prob. Chi-Square(2)	0.0008

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 10/30/19 Time: 11:32
Sample: 1990 2015
Included observations: 26
Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-11.23005	17.69095	-0.634791	0.5321
X	0.008079	0.009339	0.864992	0.3964
RESID(-1)	0.912378	0.213026	4.282949	0.0003
RESID(-2)	-0.162607	0.242997	-0.669172	0.5103

R-squared	0.545376	Mean dependent var	-2.44E-13
Adjusted R-squared	0.483382	S.D. dependent var	59.06719
S.E. of regression	42.45524	Akaike info criterion	10.47542
Sum squared resid	39653.84	Schwarz criterion	10.66897
Log likelihood	-132.1804	Hannan-Quinn criter.	10.53115
F-statistic	8.797205	Durbin-Watson stat	1.988405
Prob(F-statistic)	0.000506		

自相关的修正：广义差分法： 关键是 ρ 未知需要估计

1. 由DW=0.508796计算 $\hat{\rho}$

$$\hat{\rho} \approx 1-\frac{DW}{2} = 1-\frac{0.508796}{2} = 0.745602$$

生成广义差分变量： $Y_t^* = Y_t - 0.745602Y_{t-1}$ $X_t^* = X_t - 0.745602X_{t-1}$

输入: “ls Y-0.745602*Y(-1) C X-0.745602*X(-1)”回车

由于使用了广义差分数据，样本容量减少了1个，为25个。查5%显著水平的DW 统计表可知 dL= 1.288，dU = 1.454，模型中 DW=1.706530，d_U<DW<4-d_U 说明广义差分模型中已无自相关。

（再做个BG检验更严谨，下同！）

Dependent Variable: Y-0.745602*Y(-1)
Method: Least Squares
Date: 10/30/19 Time: 11:50
Sample (adjusted): 1991 2015
Included observations: 25 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-23.77859	15.08045	-1.576782	0.1285
X-0.745602*X(-1)	0.823006	0.023105	35.62021	0.0000

R-squared	0.982195	Mean dependent var	434.2474
Adjusted R-squared	0.981421	S.D. dependent var	289.0184
S.E. of regression	39.39432	Akaike info criterion	10.26174
Sum squared resid	35693.99	Schwarz criterion	10.35925
Log likelihood	-126.2717	Hannan-Quinn criter.	10.28878
F-statistic	1268.799	Durbin-Watson stat	1.706530
Prob(F-statistic)	0.000000		

2.德宾两步法估计ρ

第一步： 作回归 $Y_t = b_1 + \rho Y_{t-1} + b_2 X_t + b_3 X_{t-1} + \varepsilon_t$

在命令窗中输入“LS Y C Y(-1) X X(-1)”，敲回车即可。

估计结果 $\hat{Y}_t = -12.49010 + 0.883835 Y_{t-1} + 1.005760 X_t - 0.926041 X_{t-1}$

第二步： 以 $\hat{\rho} = 0.883835$ 作广义差分，生成新序列

$$Y_t^* = Y_t - 0.883835 Y_{t-1}$$
$$X_t^* = X_t - 0.883835 X_{t-1}$$

作Y_t^*与X_t^* 的回归，结果为

$$d_U < DW = 1.9312 < 4 - d_U$$

表明广义差分模型中已无自相关。

Dependent Variable: Y-0.883835*Y(-1)
Method: Least Squares
Date: 11/02/19 Time: 11:27
Sample (adjusted): 1991 2015
Included observations: 25 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-24.23733	14.04559	-1.725618	0.0978
X-0.883835*X(-1)	0.857870	0.035413	24.22474	0.0000

R-squared	0.962285	Mean dependent var	259.4168
Adjusted R-squared	0.960645	S.D. dependent var	195.5092
S.E. of regression	38.78516	Akaike info criterion	10.23057
Sum squared resid	34598.64	Schwarz criterion	10.32808
Log likelihood	-125.8821	Hannan-Quinn criter.	10.25762
F-statistic	586.8378	Durbin-Watson stat	1.931222
Prob(F-statistic)	0.000000		

3.用残差序列估计 ρ

由模型可得残差序列 e_t（注意：一定要是原方程OLS估计后的残差），使用e_t 进行滞后一期的过原点自回归，在EViews命令栏中输入ls e e (-1)可得回归方程：

$$\hat{e}_t = 0.764142 e_{t-1}$$

可知$\hat{\rho}$ =0.764142，对原模型进行广义差分：

$$Y_t^* = Y_t - 0.764142 Y_{t-1} \qquad X_t^* = X_t - 0.764142 X_{t-1}$$

作广义差分方程回归，在EViews命令栏中输入

ls Y-0.764142*Y(-1) C X-0.764142*X(-1)

回车后可得方程输出结果

广义差分输出结果

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: Y-0.764142*Y(-1)
Method: Least Squares
Date: 12/01/19 Time: 12:17
Sample (adjusted): 1991 2015
Included observations: 25 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-23.89609	14.95454	-1.597915	0.1237
X-0.764142*X(-1)	0.826050	0.024188	34.15126	0.0000

R-squared	0.980661	Mean dependent var	410.7989
Adjusted R-squared	0.979820	S.D. dependent var	276.3017
S.E. of regression	39.25020	Akaike info criterion	10.25441
Sum squared resid	35433.30	Schwarz criterion	10.35192
Log likelihood	-126.1801	Hannan-Quinn criter.	10.28145
F-statistic	1166.309	Durbin-Watson stat	1.742973
Prob(F-statistic)	0.000000		

$$d_U = 1.454 < DW = 1.742973 < 4 - d_U = 2.546$$

说明广义差分模型中已无自相关。

4.科克兰内(Cochrane)—奥克特(Orcutt)迭代法：

Eviews中命令栏输入 “LS Y C X AR(1)”/回车,即自动迭代得科克兰内-奥克特法估计结果。

由于 $d_U = 1.454 < DW = 1.966840 < 4 - d_U = 2.546$

表明已消除自相关。故可以报告其估计结果

$$\hat{Y}_t = -312.2273 + 0.877084 X_t$$
$$t = (-0.5323) \quad (9.7332)$$
$$R^2 = 0.9978 \quad F = 5016.198$$

Dependent Variable: Y
Method: Least Squares
Date: 11/02/19 Time: 11:46
Sample (adjusted): 1991 2015
Included observations: 25 after adjustments
Convergence achieved after 13 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-312.2273	586.5228	-0.532336	0.5998
X	0.877084	0.090113	9.733159	0.0000
AR(1)	0.923335	0.150782	6.123632	0.0000

R-squared	0.997812	Mean dependent var	1377.250
Adjusted R-squared	0.997613	S.D. dependent var	811.3059
S.E. of regression	39.63805	Akaike info criterion	10.30962
Sum squared resid	34565.84	Schwarz criterion	10.45589
Log likelihood	-125.8703	Hannan-Quinn criter.	10.35019
F-statistic	5016.198	Durbin-Watson stat	1.966840
Prob(F-statistic)	0.000000		

Inverted AR Roots .92

注意：迭代法结果可以直接报告，截距项不用再除。但AR(1)不能算作一个解释变量，报告回归结果时也不写，其系数其实是 ρ 的估计。

还原为原模型结果

用DW估计 ρ ： $\hat{\rho} \approx 0.7456$

$$\hat{Y}_t^* = -23.7886 + 0.8230 X_t *$$

$$\beta_1 = \beta_1^* / (1-\rho) = -23.786 / (1-0.7456) = -93.4693$$

还原为 $\hat{Y}_t = -93.4693 + 0.8230 X_t$

德宾两步法： $\hat{\rho} \approx 0.8838$

$$\hat{Y}_t^* = -24.2374 + 0.8579 X_t^*$$

$$\beta_1 = \beta_1^* / (1-\rho) = -24.2374 / (1-0.8838) = -208.5838$$

还原为 $\hat{Y}_t = -208.5838 + 0.8579 X_t$

用残差直接估计：　$\hat{\rho} = 0.7641$

$$\hat{Y}_t^* = -23.8958 + 0.8260 X_t^*$$

$$\beta_1 = \beta_1^* / (1-\rho) = -23.8958 / (1-0.7641) = -101.2963$$

还原为　$\hat{Y}_t = -101.2963 + 0.8260 X_t$

科克兰内-奥克特法：

$$\hat{Y}_t = -312.2273 + 0.8771 X_t$$

提醒1：除了科克兰内-奥克特迭代法外，其他三种方法均要记得换算回去得到原模型估计结果。Y*和X*对应的广义差分模型中R2和F统计量等不能直接使用（迭代法的可以使用）， R2也不宜相互比较。

提醒2：其实只要找到一种方法消除自相关（消除是指两种检验方法都显示无自相关才严谨）即可，不用非要比较哪种方法更好。

作业

本科教材练习题6.2（P155)