

## 第 14 章 单位根与协整

### 14.1 非平稳序列

如果时间序列不平稳，则称为非平稳序列(non-stationary time series)，主要包括以下三种情形。

(1) 确定性趋势：如果时间序列有确定性趋势(deterministic trend)，则为非平稳序列。考虑以下模型：

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t \quad (14.1)$$

其中,  $t$  为时间趋势(time trend), 而  $\beta_1 t$  为时间趋势项。

对上式两边取期望可得

$$E(y_t) = \beta_0 + \beta_1 t \quad (14.2)$$

$E(y_t)$  随时间而改变, 故不是平稳序列。

对于这种非平稳序列, 只要把时间趋势去掉, 就变成平稳序列, 故称为**趋势平稳**(trend stationary)序列。

对于趋势平稳序列, 通常的处理方法是, 直接将时间趋势( $t$ )作为解释变量放入回归方程(14.1), 然后照常使用大样本 OLS 理论进行统计推断。

(2) 结构变动(structural break): 如果一个时间序列存在结构变动(structural break), 则为非平稳序列。考虑如下模型:

$$y_t = \begin{cases} \alpha_1 + \beta_1 x_t + \varepsilon_t, & \text{若 } t < \bar{t} \\ \alpha_2 + \beta_2 x_t + \varepsilon_t, & \text{若 } t \geq \bar{t} \end{cases} \quad (14.3)$$

其中,  $\bar{t}$  为给定时间(常数)。

如果  $\alpha_1 \neq \alpha_2$  或  $\beta_1 \neq \beta_2$ , 则存在结构变动。

此时,  $E(y_t)$  在  $t = \bar{t}$  处存在跳跃, 故为非平稳序列。

对于结构变动, 可进行邹检验。

如果发现存在结构变动，可定义如下虚拟变量：

$$D_t = \begin{cases} 1, & \text{若 } t \geq \bar{t} \\ 0, & \text{其他} \end{cases} \quad (14.4)$$

然后将虚拟变量  $D_t$  引入以下回归方程：

$$y_t = \alpha_1 + \beta_1 x_t + \gamma D_t + \delta D_t x_t + \varepsilon_t \quad (14.5)$$

方程(14.5)与方程(14.3)等价；其中， $\alpha_2 = \alpha_1 + \gamma$ ，而  $\beta_2 = \beta_1 + \delta$ 。

在方程(14.5)中，所有参数都不随时间而变(就此方程而言，不再有结构变动)，故可照常进行回归。

(3) 随机趋势：另一种导致非平稳的趋势为随机趋势(stochastic trend)。比如，随机游走模型(random walk)：

$$y_t = y_{t-1} + \varepsilon_t \quad (14.6)$$

其中， $\{\varepsilon_t\}$ 为白噪声。假设时间开始于 $t = 0$ ，则

$$\begin{aligned} y_1 &= y_0 + \varepsilon_1 \\ y_2 &= y_1 + \varepsilon_2 = y_0 + \varepsilon_1 + \varepsilon_2 \\ y_3 &= y_2 + \varepsilon_3 = y_0 + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 \\ &\vdots \end{aligned} \quad (14.7)$$

$$y_t = y_{t-1} + \varepsilon_t = y_0 + \varepsilon_1 + \cdots + \varepsilon_t = y_0 + \sum_{s=1}^t \varepsilon_s$$

如果 $\varepsilon_1$ 增加一个单位,所有 $\{y_1, y_2, \dots, y_t, \dots\}$ 都将增加一个单位。

来自 $\{\varepsilon_t\}$ 的任何扰动对 $\{y_t\}$ 都具有永久的效应(permanent effect),其影响力不随时间而衰减,故称 $\{\varepsilon_t\}$ 为此模型的“随机趋势”。

在上式最后一个方程的两边,同时求方差可得

$$\text{Var}(y_t) = \text{Var}\left(\sum_{s=1}^t \varepsilon_s\right) = \sum_{s=1}^t \text{Var}(\varepsilon_s) = t\sigma_\varepsilon^2 \quad (14.8)$$

其中, $\sigma_\varepsilon^2$ 为扰动项 $\varepsilon_s$ 的方差。

当 $t \rightarrow \infty$ 时,  $\text{Var}(y_t) \rightarrow \infty$ (方差发散), 故 $\{y_t\}$ 为非平稳序列。

在方程(14.6)中, 如果包含常数项, 则为带漂移的随机游走(random walk with drift):

$$y_t = \beta_0 + y_{t-1} + \varepsilon_t \quad (14.9)$$

其中,  $\beta_0 \neq 0$  为每个时期的平均漂移(drift), 因为 $E(y_t) = \beta_0 + y_{t-1}$ 。

随机游走是 AR(1)的特例。

对于 AR(1)模型,  $y_t = \beta_0 + \beta_1 y_{t-1} + \varepsilon_t$ , 如果 $\beta_1 = 1$ , 则为随机游走。

在方程(14.9)中，将右边的  $y_{t-1}$  移项到左边可得

$$\Delta y_t = \beta_0 + \varepsilon_t \quad (14.10)$$

因此，对于随机游走，只要对其进行一阶差分，即可得到平稳序列，故称为“差分平稳”(difference stationary)序列。

**定义** 称平稳的时间序列为**零阶单整**(Integrated of order zero)，记为  $I(0)$ 。如果时间序列的一阶差分为平稳过程，则称为**一阶单整**(Integrated of order one)，记为  $I(1)$ ，也称为**单位根过程**(unit root process)。更一般地，如果时间序列的  $d$  阶差分为平稳过程，则称为  $d$ **阶单整**(Integrated of order  $d$ )，记为  $I(d)$ 。



对于  $I(0)$  序列，由于它是平稳的，故长期而言有回到其期望值的趋势。

这种性质被称为均值回复(mean-reverting)。

非平稳的  $I(1)$  序列则会“到处乱跑”(wander widely)，没有上述性质。

比如，随机游走的方差越来越大，趋向无穷。

$I(0)$  序列对于其过去的行为只有有限的记忆，即发生在过去的扰动项对未来的影响随时间而衰减。

$I(1)$  序列则对过去的行为具有无限长的记忆，即任何过去的冲击都将永久性地改变未来的整个序列。

例 如果 $\{GDP_t\}$ 为  $I(1)$ , 则任何货币政策或财政政策的调整都将对未来 GDP 产生永久影响。

定义 如果时间序列 $\{y_t\}$ 的  $d$  阶差分为平稳的  $ARMA(p, q)$ 过程, 则称 $\{y_t\}$ 为  $ARIMA(p, d, q)$ 过程。

最常见的为  $ARIMA(p, 1, q)$ , 即经过一次差分就得到平稳的  $ARMA(p, q)$ 。

## 14.2 ARMA 的平稳性

在什么情况下,  $ARMA(p, q)$ 才平稳呢?

$MA(q)$ 是平稳的, 因为它是有限个白噪声的线性组合。

ARMA( $p, q$ )的平稳性仅取决于 AR( $p$ )的部分。

首先, 考虑 AR(1)模型:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \varepsilon_t \quad (14.11)$$

如果 $|\beta_1| < 1$ , 则 $\{y_t\}$ 为平稳过程。

上式其实是一阶随机差分方程, 其稳定性与对应的以下确定性差分方程是一样的:

$$y_t = \beta_0 + \beta_1 y_{t-1} \quad (14.12)$$

只要考虑一阶差分方程(14.12)是否有稳定解即可。

方程(14.12)是一个非齐次的差分方程(含常数项 $\beta_0$ ), 它的解又取决于相应的齐次差分方程(不含常数项):

$$y_t = \beta_1 y_{t-1} \quad (14.13)$$

此齐次差分方程的通解为

$$y_t = y_0 \beta_1^t \quad (14.14)$$

由于此通解的形式为指数函数, 故 AR(1)模型(14.11)的稳定性条件为 $|\beta_1| < 1$ 。

考虑 AR( $p$ )模型:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + \varepsilon_t \quad (14.15)$$

此 AR( $p$ )模型的稳定性条件取决于相应的确定性齐次差分方程:

$$y_t = \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} \quad (14.16)$$

仍然假设差分方程(14.16)的解形式为指数函数, 即

$$y_t = z^{-t} = (1/z)^t \quad (14.17)$$

其中,  $z$ 的取值待定。将表达式(14.17)代入差分方程(14.16)可得

$$z^{-t} - \beta_1 z^{-(t-1)} - \cdots - \beta_p z^{-(t-p)} = 0 \quad (14.18)$$

上式两边同乘以  $z^t$  可得 AR( $p$ )模型的特征方程(characteristic equation):

$$\phi(z) \equiv 1 - \beta_1 z - \cdots - \beta_p z^p = 0 \quad (14.19)$$

此  $p$  阶多项式方程(14.19)在复数域中一定有  $p$  个根(包括重根), 比如  $(z_1, z_2, \cdots, z_p)$ 。

与此对应, 齐次差分方程(14.16)也有  $p$  个形如  $(1/z)^t$  的解, 而其通解则是这  $p$  个解的线性组合:

$$y_t = k_0 + k_1 (1/z_1)^t + k_2 (1/z_2)^t + \cdots + k_p (1/z_p)^t \quad (14.20)$$

其中,  $(k_0, k_1, \cdots, k_p)$  为待定常数, 取决于初始条件  $\{y_0, y_1, \cdots, y_{p-1}\}$ 。

如果要求  $\{y_t\}$  收敛于一个稳定值，则在方程(14.20)中，所有  $(1/z_j)^t$  ( $j = 1, \dots, p$ ) 均应收敛到 0。

由于  $z_j$  为复数，这意味着，特征方程所有解的范数  $\|z_j\|$  (即在复平面上  $z_j$  离原点的距离) 都必须大于 1。特征方程的所有解必须都落在复平面上的单位圆之外(参见图 14.1)。

如果特征方程的某个根落在单位圆之内，则为爆炸式增长的非平稳过程。

如果某个根正好落在单位圆之上，则称为单位根(unit root)，比如随机游走的情形。

例 对于 AR(1), 其特征方程为  $1 - \beta_1 z = 0$ , 故  $z = 1/\beta_1$ 。因此,  $\|z\| = |z| > 1 \Leftrightarrow |\beta_1| < 1$ 。由此可见, 有关 AR( $p$ ) 平稳性的结论是对 AR(1) 情形的推广。

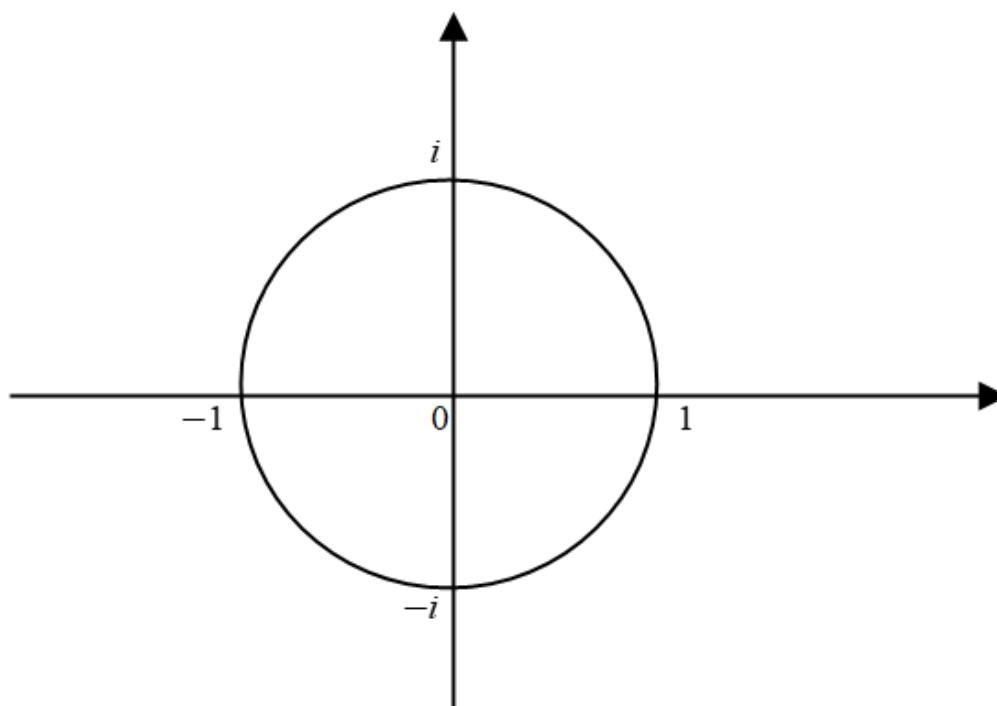


图 14.1 复平面上的单位圆



### 14.3 VAR 的平稳性

AR( $p$ )的平稳性条件可推广到多维 VAR( $p$ )的情形。

考虑以下 VAR( $p$ )模型：

$$\mathbf{y}_t = \boldsymbol{\Gamma}_0 + \boldsymbol{\Gamma}_1 \mathbf{y}_{t-1} + \cdots + \boldsymbol{\Gamma}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t \quad (14.21)$$

其中， $\boldsymbol{\varepsilon}_t$ 为向量白噪声过程。如果特征方程

$$\left| \mathbf{I}_n - \boldsymbol{\Gamma}_1 z - \cdots - \boldsymbol{\Gamma}_p z^p \right| = 0 \quad (14.22)$$

的所有根都落在复平面的单位圆之外(即 $\|z\| > 1$ )，则此 VAR( $p$ )为平稳过程，其中 $|\cdot|$ 表示行列式。

此平稳性的等价条件为伴随矩阵(companion matrix)

$$\tilde{\Gamma} = \begin{pmatrix} \Gamma_1 & \Gamma_2 & \cdots & \Gamma_p \\ \mathbf{I}_n & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_n & \mathbf{0} \end{pmatrix}_{np \times np} \quad (14.23)$$

的所有特征值都落在单位圆之内。

## 14.4 单位根所带来的问题

对于 AR(1)模型,一般从理论上认为,不太可能出现自回归系数  $|\beta_1| > 1$  的情形; 否则任何对经济的扰动都将被无限放大。

经济学家通常只担心存在单位根的情形，即  $\beta_1 = 1$ 。

如果时间序列存在单位根，则非平稳，可能带来以下问题。

(1) 自回归系数的估计值不服从渐近正态分布， $t$ 检验失效

(2)

考虑以下 AR(1)模型：

$$y_t = \beta_0 + \beta_1 y_{t-1} + \varepsilon_t \quad (14.24)$$

假设真实值  $\beta_1 = 1$ ，故为单位根过程。

对上式进行 OLS 回归，可得  $\beta_1$  的 OLS 估计量  $\hat{\beta}_1$ 。

由于存在单位根( $\beta_1$ 的真实值为 1),  $\hat{\beta}_1$ 不服从渐近正态分布, 甚至不是对称分布(即使是在大样本中), 而是向左偏向于 0(分布左边有很长的尾巴)。

由于 $\{y_t\}$ 不是平稳序列, 故中心极限定理不再适用。

虽然 $\text{plim}_{n \rightarrow \infty} \hat{\beta}_1 = \beta_1$ (仍为一致估计), 但在有限样本下可能存在较大偏差。

由于 $\hat{\beta}_1$ 不是渐近正态分布, 故 $t$ 统计量也不服从渐近标准正态分布, 无法进行传统的区间估计与假设检验。

建立于平稳性假设基础之上的大样本理论不再适用。

下面通过蒙特卡罗法考察  $\hat{\beta}_1$  的大样本分布。考虑不带漂移项的随机游走：

$$y_t = y_{t-1} + \varepsilon_t \quad (14.25)$$

其中，扰动项  $\varepsilon_t$  为 iid，且服从标准正态分布。

假设  $y_0 = 0$ ，则  $y_1 = \varepsilon_1$ ， $y_2 = \varepsilon_1 + \varepsilon_2$ ， $\dots$ ，  
 $y_t = \varepsilon_1 + \dots + \varepsilon_t = \sum_{s=1}^t \varepsilon_s$ 。

假设样本容量为 1000，首先从标准正态分布随机抽取 1000 个扰动项的观测值  $\{\varepsilon_1, \dots, \varepsilon_{1000}\}$ ，由此生成  $y_t$  的 1000 个观测值  $\{y_1, \dots, y_{1000}\}$ ，然后根据方程(14.24)进行 OLS 回归，得到自回归系数估计值  $\hat{\beta}_1$ 。

重复此过程 1,000 次，即可得到 1,000 个  $\hat{\beta}_1$ ，由此获得  $\hat{\beta}_1$  的大样本分布。

在 Stata 中首先定义一个叫 “randwalk” 的程序来产生随机游走，进行一阶自回归，并得到一个自回归系数  $\hat{\beta}_1$ ：

```
program randwalk, rclass      (定义程序 “randwalk”，
                               并以 r() 形式储存结果)
    drop _all                 (删去内存中已有数据)
    set obs 1000               (确定样本容量为 1000)
    gen eps=rnormal()          (产生服从标准正态分布的扰动项  $\varepsilon_t$ )
    gen y=sum(eps)             (假设  $y_0 = 0$ ，定义随机游走  $y_t = \sum_{s=1}^t \varepsilon_s$ )
    gen t=_n                   (定义时间变量，第  $t$  期即第  $i$  个观测值)
```

```

tsset t      (将数据设为时间序列，以便使用滞后算子)
reg y L.y    (回归  $y_t = \beta_0 + \beta_1 y_{t-1} + error$ )
return scalar b1=_b[L.y]  (记 OLS 系数  $\hat{\beta}_1$  为 b1)
end          (程序结束)

```

然后，使用命令“simulate”来执行“randwalk”程序 1,000 遍，得到  $\hat{\beta}_1$  的大样本分布，并画其经验的概率密度图(结果参见图 14.2)。

```

. simulate beta=r(b1),seed(10101) reps(1000):
randwalk
. kdensity beta

```

命令“kdensity”表示 kernel density，即画核密度图。

直方图为对概率密度函数的不连续估计，而核密度图则为对概率密度函数的连续估计。

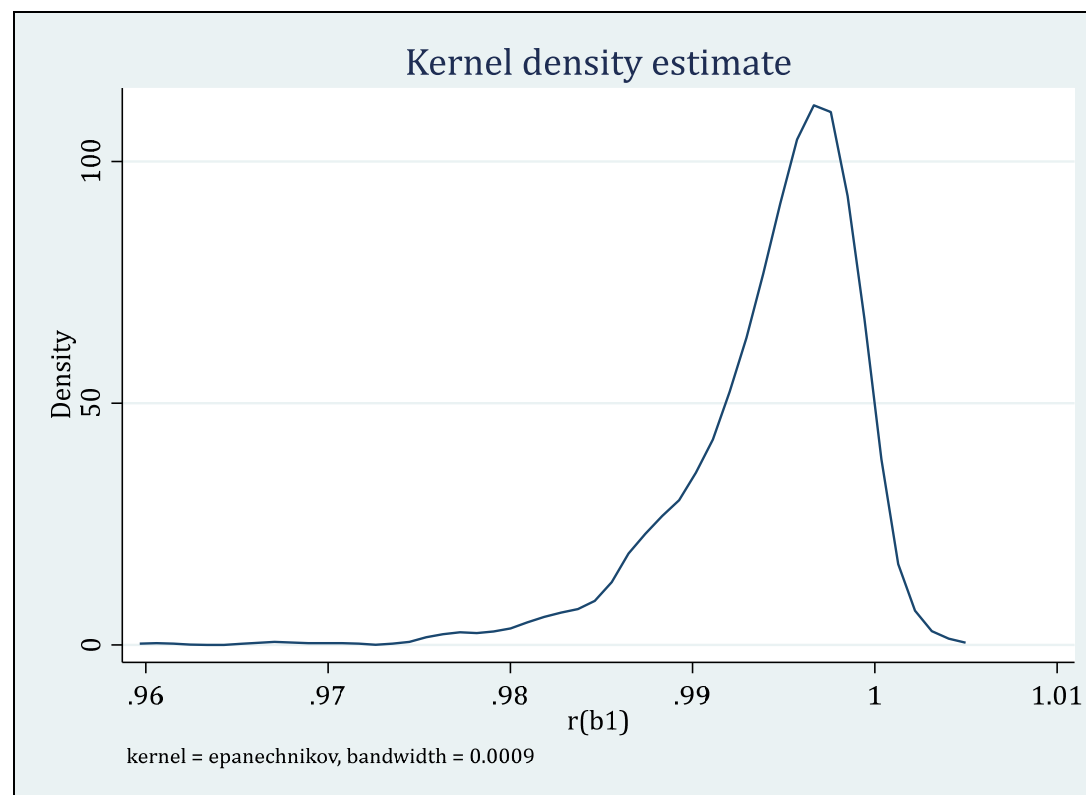


图 14.2 在单位根情况下  $\hat{\beta}_1$  的大样本分布

即使在样本容量为 1000 的情况下，OLS 估计量  $\hat{\beta}_1$  的分布也不对称，向左偏向于 0，与对称的正态分布相去甚远。



## (2) 两个相互独立的单位根变量可能出现伪相关或伪回归

单位根变量还可能出现的另一严重后果是，即使两个单位根变量相互独立，根据样本数据进行相关分析或回归分析，却可能发现二者有显著的关系，这被称为**伪相关**(spurious correlation)或**伪回归**(spurious regression)。

考虑以下两个单位根过程：

$$y_t = y_{t-1} + u_t; \quad x_t = x_{t-1} + v_t \quad (14.26)$$

其中， $u_t, v_t$ 均为 iid 且相互独立。

因此， $y_t$ 与 $x_t$ 也是相互独立的。

考虑以下 OLS 回归：

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad (14.27)$$

由于  $y_t$  与  $x_t$  相互独立，故在上式中真实参数  $\beta = 0$ 。

如果样本容量足够大，则期待 OLS 估计值  $\hat{\beta} \approx 0$ ， $R^2 \approx 0$ ，但实际结果并非如此，因为扰动项  $\varepsilon_t = y_t - \alpha - \beta x_t$  非平稳。

这一现象最初由 Granger and Newbold (1974) 通过蒙特卡罗模拟发现。

下面在 Stata 中模拟此过程。

假设  $y_0 = 0$ ， $x_0 = 0$ ，则  $y_t = \sum_{s=1}^t u_s$ ， $x_t = \sum_{s=1}^t v_s$ 。

假设样本容量为 10,000，首先在 Stata 中生成相互独立的单位根变量  $y_t$  与  $x_t$ ，然后进行 OLS 回归。

. drop _all	(删去内存中已有数据)
. set obs 10000	(确定样本容量为 10,000)
. set seed 1234	(确定随机数的种子为 1234)
. gen u=rnormal( )	(产生服从标准正态分布的扰动项 $u_t$ )
. gen y=sum(u)	(定义随机游走 $y_t = \sum_{s=1}^t u_s$ )
. set seed 12345	(确定随机数的种子为 12345)
. gen v=rnormal( )	(产生服从标准正态分布的扰动项 $v_t$ )
. gen x=sum(v)	(定义随机游走 $x_t = \sum_{s=1}^t v_s$ )
. reg y x	

Source	SS	df	MS	Number of obs	=	10,000
				F(1, 9998)	=	866.95
Model	2406934.55	1	2406934.55	Prob > F	=	0.0000
Residual	27757621.9	9,998	2776.31745	R-squared	=	0.0798
				Adj R-squared	=	0.0797
Total	30164556.4	9,999	3016.75732	Root MSE	=	52.691

y	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
x	-.4000392	.0135864	-29.44	0.000	-.4266713	-.3734071
_cons	-55.43203	1.104488	-50.19	0.000	-57.59705	-53.26701

尽管  $y_t$  与  $x_t$  相互独立(因为使用了不同的随机数种子, 故  $u_t$  与  $v_t$  相互独立), 但  $y_t$  对  $x_t$  的回归系数却在 1% 水平上显著( $t$  值为 -29.44), 而且  $R^2$  达到 0.08, 故存在“伪回归”。

进一步, 把  $u_t$  与  $v_t$  回归。

```
. reg u v
```

Source	SS	df	MS	Number of obs	=	10,000
				F(1, 9998)	=	0.49
Model	.500659016	1	.500659016	Prob > F	=	0.4819
Residual	10121.5297	9,998	1.01235544	R-squared	=	0.0000
				Adj R-squared	=	-0.0001
Total	10122.0304	9,999	1.01230427	Root MSE	=	1.0062
u	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
v	-.0070983	.0100937	-0.70	0.482	-.0268841	.0126874
_cons	-.009079	.0100624	-0.90	0.367	-.0288033	.0106452

$u_t$ 对 $v_t$ 的回归系数高度不显著( $p$ 值高达 0.482,  $R^2$ 为 0.0000); 但由 $u_t$ 与 $v_t$ 所产生的随机游走过程 $y_t$ 与 $x_t$ 却显著相关。

看一下 $y_t$ 与 $x_t$ 的时间趋势图, 结果参见图 14.3。

```
. gen t=_n      (定义时间变量t)
. line y x t,lp(dash)
```

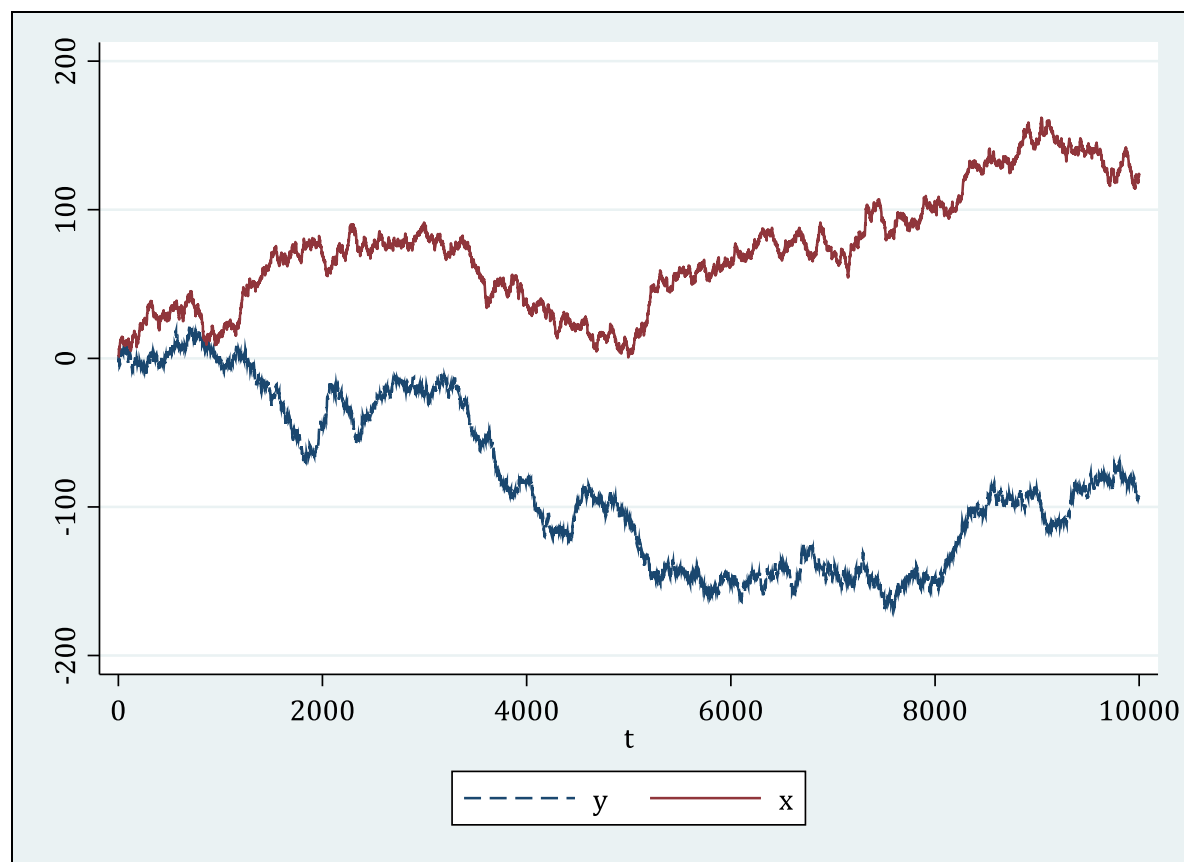


图 14.3 伪相关示意图

尽管  $y_t$  与  $x_t$  相互独立，但因为二者都是单位根过程，故存在相似的时间趋势而出现伪相关与伪回归，从而误导统计推断。

如何避免伪相关或伪回归？

方法之一，先对  $I(1)$  变量作一阶差分，得到平稳的  $I(0)$  序列，再作回归。

方法之二为“协整” (cointegration)。

但首先必须检验是否存在单位根。

## 14.5 单位根检验

### 1. Dickey-Fuller 单位根检验(Dickey and Fuller, 1979)

考虑如下  $AR(1)$  模型：

$$y_t = \beta_0 + \beta_1 y_{t-1} + \gamma t + \varepsilon_t \quad (14.28)$$

其中， $\varepsilon_t$ 为白噪声， $\gamma t$ 为时间趋势(如不含时间趋势，则 $\gamma = 0$ )。

考虑以下单边检验：

$$H_0 : \beta_1 = 1 \quad \text{vs} \quad H_1 : \beta_1 < 1 \quad (14.29)$$

其中，替代假设为“ $H_1 : \beta_1 < 1$ ”，因为理论上认为不可能出现 $\beta_1 > 1$ (爆炸式增长)的情形。

如果 $H_0$ 成立，则 $y_t$ 为带漂移项 $\beta_0$ 的随机游走；如果不带漂移项，可令 $\beta_0 = 0$ 。

但我们通常希望检验参数是否等于 0。



在原方程两边同时减去  $y_{t-1}$  可得

$$\Delta y_t = \beta_0 + \delta y_{t-1} + \gamma t + \varepsilon_t \quad (14.30)$$

其中,  $\delta \equiv \beta_1 - 1$ 。对应的原假设与替代假设变为

$$H_0 : \delta = 0 \quad vs \quad H_1 : \delta < 0 \quad (14.31)$$

对方程(14.30)作 OLS 回归, 可得估计量  $\hat{\delta}$  及相应  $t$  统计量(使用普通标准误)。

此  $t$  统计量称为 DF 统计量(Dickey-Fuller statistic, 简记 DF), 在 Stata 中记为  $Z(t)$ 。

但 $Z(t)$ 并不服从渐近正态分布，其临界值须通过蒙特卡罗模拟获得。

$Z(t)$ 越小(绝对值很大的负数)，则越倾向于拒绝原假设。

故 DF 检验是左边单侧检验，其拒绝域只在分布的最左边。

## 2 . Augmented Dickey-Fuller 单位根检验 (Dickey and Fuller, 1981)

DF 检验使用一阶自回归来检验单位根，但要求扰动项 $\{\varepsilon_t\}$ 为白噪声，故扰动项无自相关。

如果 $\{\varepsilon_t\}$ 存在自相关, 可在方程(14.30)中以引入被解释变量 $\Delta y_t$ 的高阶滞后项(即滞后差分项), 以保证扰动项 $\{\varepsilon_t\}$ 为白噪声:

$$\Delta y_t = \beta_0 + \delta y_{t-1} + \gamma_1 \Delta y_{t-1} + \gamma_2 \Delta y_{t-2} + \cdots + \gamma_{p-1} \Delta y_{t-p+1} + \gamma t + \varepsilon_t \quad (14.32)$$

而原假设与替代假设依然为

$$H_0 : \delta = 0 \quad vs \quad H_1 : \delta < 0 \quad (14.33)$$

对此方程使用 OLS 可得估计量 $\hat{\delta}$ 及相应的 $t$ 统计量(使用普通标准误)。

此 $t$ 统计量称为增广 DF 统计量 (Augmented Dickey-Fuller statistic, 简记 ADF), Stata 仍记其为 $Z(t)$ 。

ADF 检验是最常用的单位根检验。

与 DF 检验一样，ADF 检验也是左边单侧检验，其拒绝域只在分布的最左边。

ADF 统计量的临界值也要通过蒙特卡罗模拟得到。

ADF 统计量的临界值取决于真实模型( $H_0$ )是否带漂移项，以及用于检验的回归方程(14.32)是否包含常数项或时间趋势。

Stata 手册(Stata Manual)总结了以下四种情形，参见表 14.1。

表 14.1 ADF 检验的四种情形

情形	真实模型( $H_0$ )	对回归方程(14.32)的约束	Stata 选择项
1	不带漂移项	$\beta_0 = 0, \gamma = 0$	noconstant
2	不带漂移项	$\gamma = 0$	默认值
3	带漂移项	$\gamma = 0$	drift
4	不带或带漂移项	无约束	trend

对于表 14.1 中的情形 2，虽然真实模型不包含漂移项(即无常数项)，但在 ADF 检验的回归方程中依然包括了常数项。

情形 2 与情形 3 对回归方程的约束相同，故检验统计量也相同，但临界值不同(因为真实模型不同)。

ADF 检验的 Stata 命令为

```
. dfuller y,lags(p) regress noconstant drift  
trend
```

选择项 “lags( $p$ )” 表示包含  $p$  阶滞后差分项，默认为 “lags(0)”，对应于 DF 检验。

选择项 “regress” 表示显示回归结果。

选择项 “noconstant drift trend” (三者最多选一项，不能并用)的含义参见表 14.1。

## 关于常数项与时间趋势项

在作 ADF 检验时，是否应该带常数项或时间趋势项，首先应从理论上考虑。

比如，考察 GDP 对数是否有单位根，一般应包含时间趋势项；而利率、汇率等则不应有时间趋势项。

也可通过画时间序列图来大致判断有无长期趋势。

在作 ADF 检验时，使用选择项 “regress”，可以看到常数项或时间趋势项是否显著。

如果无从判断，为了稳健起见，可以把各种情况都进行检验。

并将结果以  $(c, t, p)$  格式列表，其中 “ $c = 1$ ” 表示带常数项，“ $c = 0$ ” 表示不带常数项；“ $t = 1$ ” 表示带趋势项，“ $t = 0$ ” 表示不带趋势项；而  $p$  表示滞后阶数。

### 关于滞后阶数 $p$ 的确定

在进行 ADF 检验时，如何确定滞后阶数  $p$  是个实际问题。

ADF 检验的结果常常对滞后阶数  $p$  很敏感。

如果  $p$  太小，则扰动项  $\{\varepsilon_t\}$  可能存在自相关，使得检验出现偏差。

如果  $p$  太大，则会降低检验的功效(power)。

Schwert (1989) 建议，取最大滞后阶数为



$$p_{\max} = [12 \cdot (T/100)^{1/4}] \quad (14.34)$$

其中， $T$ 为样本容量， $[\cdot]$ 表示取整数部分；然后使用由大到小的序贯 $t$ 规则，看 ADF 检验中最后一阶回归系数是否显著。

也可使用信息准则，比如 AIC 或 BIC。

### 3. 单整阶数(order of integration)的确定

对时间序列 $\{y_t\}$ 进行单位根检验后，如果认为 $\{y_t\}$ 为非平稳，则要进一步判断其为 I(1)或 I(2)。

此时，可对一阶差分 $\{\Delta y_t\}$ 进行单位根检验，如果 $\{\Delta y_t\}$ 为平稳，则 $\{y_t\}$ 是 I(1)。

否则，要继续对二阶差分 $\{\Delta^2 y_t \equiv \Delta y_t - \Delta y_{t-1}\}$ 进行单位根检验。

如果 $\{\Delta^2 y_t\}$ 为平稳，则 $\{y_t\}$ 为 I(2)，以此类推。

在经济变量中，I(0)与 I(1)最为常见，而 I(2)很少见。

## 14.6 单位根检验的 Stata 实例

例 Nelson and Plosser (1982)使用 ADF 检验考察 1860-1970 年，美国 14 个年度宏观时间序列，结果发现只有其中一个变量可拒绝单位根的原假设，而其余 13 个变量均可视为单位根变量。此文一出，引起经济学界对单位根的广泛关注。

以数据集 `nelson_plosser.dta` 为例，包括美国的年度宏观经济变量，取自 Nelson and Plosser (1982)。

检验其中的两个变量，即 *lrgnp* (实际 GNP 对数)与 *lun* (失业率对数)，是否含有单位根。

首先，看一下这两个变量的时间趋势图(结果参见图 14.4):

```
. use nelson_plosser.dta,clear  
. tsline lrgnp lun if year>=1890,lp(dash)  
xlabel(1890(10)1970)
```

其中，变量 *lrgnp* 的取值始于 1909 年，而变量 *lun* 的取值始于 1890 年，故使用条件语句 “`if year>=1890`” 来限制画图的范围(使之更加美观)。选择项 “`xlabel(1890(10)1970)`” 表示，在横轴从 1890 年至 1970 年，每 10 年做个标注(label)。

---

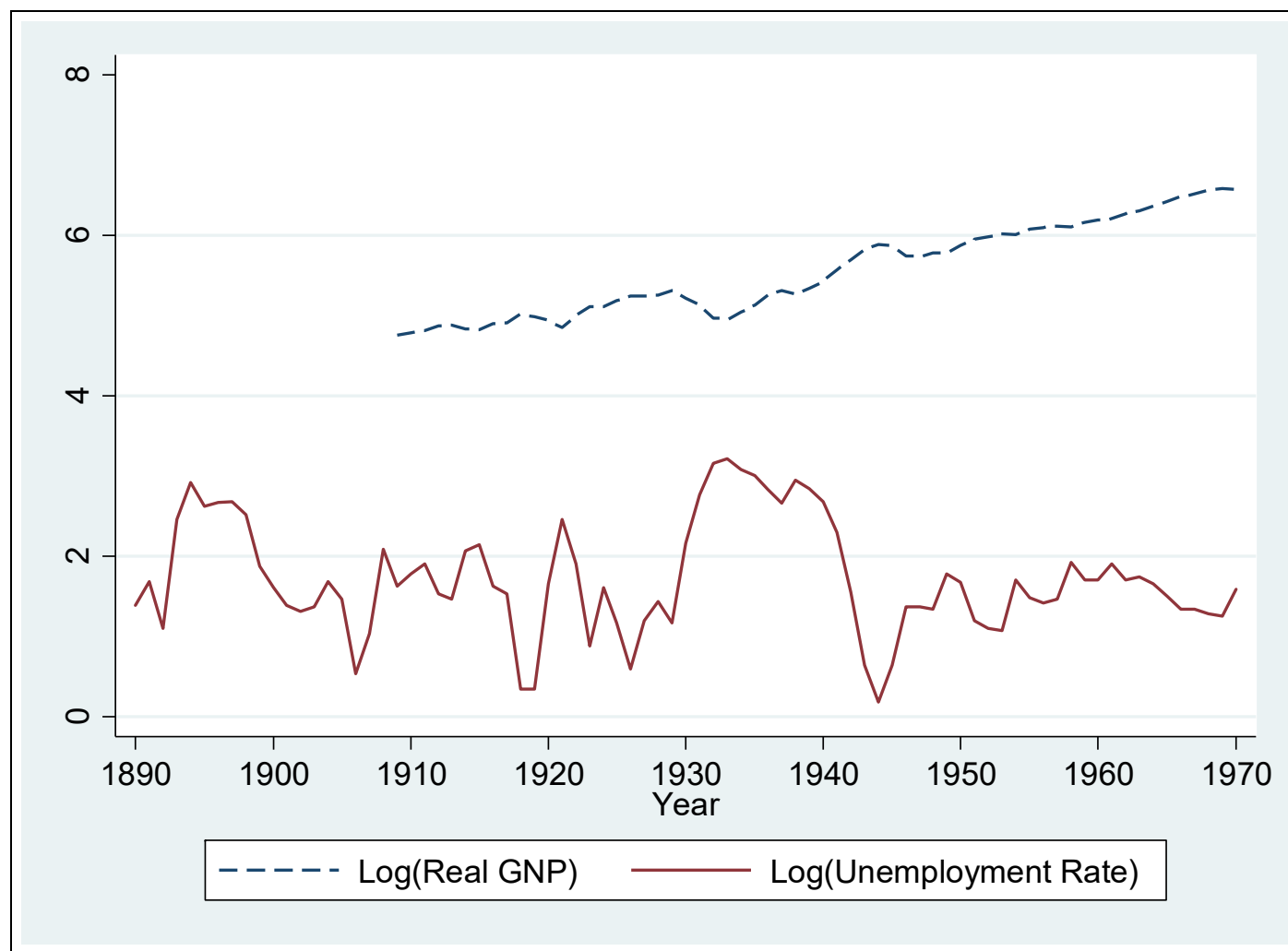


图 14.4 实际 GNP 对数与失业率对数的时间趋势

实际 GNP 对数(以虚线表示)有明显的上升时间趋势；而且较为光滑，这意味着当期值强烈地依赖于上期值，即自回归系数接近于 1，故可能为单位根过程。

失业率对数则看不出有什么时间趋势，而且较不光滑，即自回归系数明显小于 1，故不太可能为单位根过程。

对于实际 GNP 对数，首先考虑带常数项与时间趋势项的 DF 检验：

```
. dfuller lrgnp,trend
```

Dickey-Fuller test for unit root		Number of obs = 61		
Variable: lrgnp		Number of lags = 0		
H0: Random walk with or without drift				
	Test	Dickey-Fuller		
	statistic	1%	5%	10%
Z(t)	-2.026	-4.126	-3.489	-3.173
MacKinnon approximate p-value for Z(t) = 0.5871.				

DF 检验的 5%临界值为-3.489。由于 DF 统计量  $Z(t)$  为  $-2.026 > -3.489$  (左 边单侧检验), 故可在 5%的水平上接受 “存在单位根” 的原假设。

由于扰动项可能存在自相关, 考虑更高阶的 ADF 检验。

首先，计算 Schwert (1989) 建议的最大滞后阶数  $p_{\max} = [12 \cdot (T/100)^{1/4}]$ :

```
. dis 12*(62/100)^(1/4)  
10.648273
```

其中，样本容量为 62。

这表明， $p_{\max} = 10$ 。

下面，令  $\hat{p} = 10$ ，进行 ADF 检验：

```
. dfuller lrgnp, lags(9) trend reg
```

# Augmented Dickey-Fuller test for unit root

Variable: lrgnp

Number of obs = 52

Number of lags = 9

H0: Random walk with or without drift

Test statistic	Dickey-Fuller critical value			
	1%	5%	10%	
Z(t)	-2.308	-4.146	-3.498	-3.179

MacKinnon approximate p-value for Z(t) = 0.4296.

## Regression table

D.lrgnp	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
lrgnp						
L1.	-.2431874	.1053744	-2.31	0.026	-.456157	-.0302179
LD.	.4268154	.1533028	2.78	0.008	.1169789	.736652
L2D.	.1299825	.1591474	0.82	0.419	-.1916663	.4516314
L3D.	.0148792	.1601745	0.09	0.926	-.3088455	.3386039
L4D.	.0074853	.1591925	0.05	0.963	-.3142547	.3292252
L5D.	-.0628515	.1517327	-0.41	0.681	-.3695147	.2438117
L6D.	.1135881	.1488912	0.76	0.450	-.1873321	.4145084
L7D.	.1157849	.1470971	0.79	0.436	-.1815095	.4130792
L8D.	-.0360994	.1474121	-0.24	0.808	-.3340302	.2618315
L9D.	-.1339387	.1441335	-0.93	0.358	-.4252433	.157366
_trend	.0083664	.003355	2.49	0.017	.0015856	.0151471
_cons	1.091761	.4663261	2.34	0.024	.1492807	2.034241



时间趋势项(\_trend)很显著( $p$  值为 0.017), 但最后一阶滞后项(L9D.)在 5%的水平上并不显著( $p$  值为 0.358)。

依次令  $\hat{p} = 9, \dots, 3$ , 进行 ADF 检验, 最后一阶滞后项仍不显著(过程略)。

下面, 令  $\hat{p} = 2$ , 再进行 ADF 检验。

```
. dfuller lrgnp, lags(1) trend reg
```

# Augmented Dickey-Fuller test for unit root

Variable: lrgnp

Number of obs = 60

Number of lags = 1

H0: Random walk with or without drift

Test statistic	Dickey-Fuller critical value			
	1%	5%	10%	
Z(t)	-2.994	-4.128	-3.490	-3.174

MacKinnon approximate *p*-value for Z(t) = 0.1338.

## Regression table

D.lrgnp	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
lrgnp						
L1.	-.1753423	.0585665	-2.99	0.004	-.2926651	-.0580195
LD.	.4188865	.1209448	3.46	0.001	.1766048	.6611681
_trend	.0056465	.0018615	3.03	0.004	.0019174	.0093757
_cons	.8134145	.2679888	3.04	0.004	.2765684	1.350261

最后一阶滞后项(LD.)在 1%的水平上显著地不等于 0。

ADF 统计量  $z(t)$  为  $-2.994 > -3.490$ , 故无法在 5%的水平上拒绝存在单位根的原假设, 即可认为实际 GNP 对数  $lrgnp$  含有单位根。

“麦金农的近似  $p$  值” (MacKinnon approximate  $p$ -value) 为 0.1338, 与此结论一致。

进一步检验是否  $lrgnp$  的一阶差分为平稳过程。根据序贯  $t$  规则, 选择  $\hat{p} = 1$ , 进行 DF 检验。

由于  $\Delta lrgnp$  已不存在时间趋势(读者可自行画图考察), 故检验时不带时间趋势项。

```
. dfuller d.lrgnp
```

Dickey-Fuller test for unit root		Number of obs = 60		
Variable: D.lrgnp		Number of lags = 0		
H0: Random walk without drift, d = 0				
	Test	Dickey-Fuller		
	statistic	1%	5%	10%
Z(t)	-5.322	-3.566	-2.922	-2.596
MacKinnon approximate p-value for Z(t) = 0.0000.				

ADF 统计量  $Z(t)$  为  $-5.322 < -3.566$ , 故可在 1% 的水平上拒绝存在单位根的原假设, 即认为  $\Delta \text{lrgnp}$  为平稳过程。由此可知,  $\text{lrgnp}$  为  $I(1)$  过程。

下面考察失业率对数  $lun$  是否含有单位根。

根据序贯 $t$ 规则，选择  $\hat{p} = 4$ ，进行不带时间趋势的 ADF 检验：

```
. dfuller lun,lags(3) reg
```

# Augmented Dickey-Fuller test for unit root

Variable: lun

Number of obs = 77

Number of lags = 3

H0: Random walk without drift, d = 0

Test statistic	Dickey-Fuller critical value			
	1%	5%	10%	
Z(t)	-3.588	-3.542	-2.908	-2.589

MacKinnon approximate p-value for Z(t) = 0.0060.

## Regression table

D.lun	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
lun						
L1.	-.2908167	.0810476	-3.59	0.001	-.4523821	-.1292513
LD.	.3824999	.1099181	3.48	0.001	.1633822	.6016177
L2D.	-.2030521	.1033701	-1.96	0.053	-.4091166	.0030123
L3D.	.2431095	.1041622	2.33	0.022	.0354659	.4507531
_cons	.4898167	.1476383	3.32	0.001	.1955051	.7841283

ADF 统计量  $z(t)$  为  $-3.588 < -3.542$ ，故可在 1% 的水平上拒绝存在单位根的原假设，即认为失业率对数  $lun$  为平稳过程。麦金农近似  $p$  值为 0.0060，与此结论一致。

## 14.7 协整的思想与初步检验

对于有单位根的变量，传统的处理方法是先进行一阶差分得到平稳序列，然后再对平稳序列建模。

但一阶差分后变量的经济含义与原序列并不相同，而有时仍希望用原序列进行回归。

如果多个单位根变量之间由于某种经济力量而存在长期均衡关系(long-run equilibrium)，则有可能进行这种回归。

其基本思想是,如果多个单位根序列拥有**共同随机趋势**(common stochastic trend),则可以对这些变量作适当的线性组合而消去此随机趋势,从而得到平稳序列。

**例** 短期利率与长期利率可能都是单位根过程,而且二者的走势也很相似。

从经济理论上来看,长期利率是未来预期短期利率的平均值与“**风险溢价**”(risk premium)之和(持有长期资产面临更大风险),故存在长期均衡关系。



以数据集 `macro_3e.dta` 为例，该数据集包含美国 1957q1-2005q1 的季度宏观变量，时间变量为 *time*。

打开此数据集后，考察 *fygm3* (3 个月国库券利率)与 *fygt1* (1 年期国库券利率)的时间趋势，结果参见图 14.5。

```
. use macro_3e.dta,clear  
  
. tsset time  
  
. tsline fygm3 fygt1,lp(dash)
```

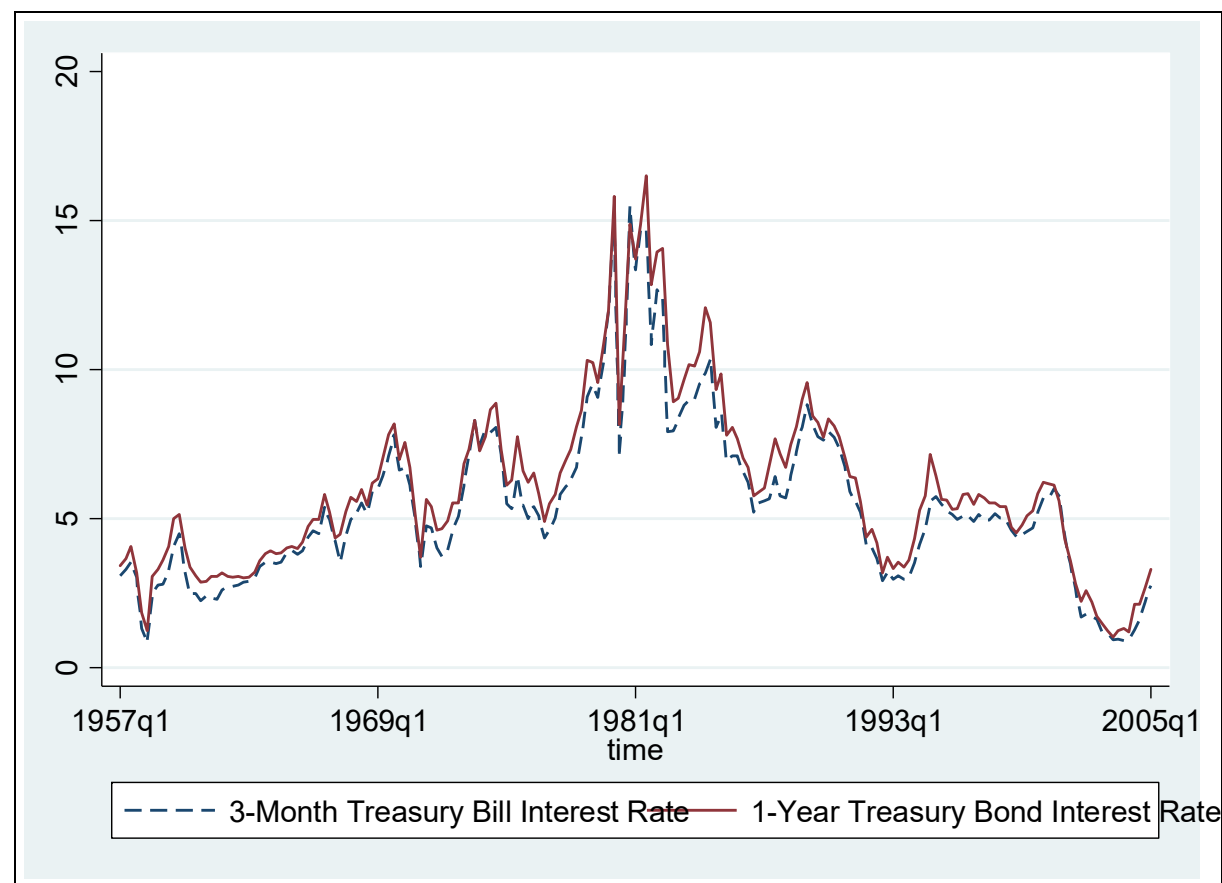


图 14.5 美国 3 月与 1 年期国库券利率的时间趋势

3 个月国库券利率(虚线)与 1 年期国库券利率(实线)的时间趋势十分接近，故二者很可能存在协整关系。

例(非正式) 当你遛狗时, 假设你与狗的每一步位置为随机游走(带漂移项), 均为单位根过程。由于你与狗之间有一根皮带相连(“长期均衡关系”), 故你与狗的位置不会相离太远(尽管二者都是单位根过程)。

假设两个  $I(1)$  过程  $\{y_t\}, \{x_t\}$  可以分别表示为

$$\begin{cases} y_t = \alpha + \beta w_t + \varepsilon_t \\ x_t = \gamma + \delta w_t + u_t \end{cases} \quad (14.35)$$

其中,  $w_t$  为随机游走,  $w_t = w_{t-1} + v_t$ ; 而  $\varepsilon_t, u_t, v_t$  均为白噪声。

由于  $\{y_t\}$  与  $\{x_t\}$  拥有共同的随机趋势  $w_t$ , 故二者的如下线性组合为平稳过程:

$$\delta y_t - \beta x_t = (\alpha\delta - \beta\gamma) + (\delta\varepsilon_t - \beta u_t) \quad (14.36)$$

其中,  $(\alpha\delta - \beta\gamma)$  为常数, 而  $(\delta\varepsilon_t - \beta u_t)$  为白噪声的线性组合。

称  $\{y_t\}$  与  $\{x_t\}$  协整(cointegrated), 称向量  $(\delta, -\beta)$  为协整向量(cointegrating vector)或协整系数。

可把协整向量  $(\delta, -\beta)$  标准化为  $(1, -\beta/\delta)$ 。

对于两个 I(1) 变量, 只可能存在一个协整关系。

对于  $n$  个 I(1) 变量, 最多可能存在  $(n-1)$  个协整关系。

一组 I(1) 变量之间协整关系的个数被称为协整秩(cointegration rank), 即线性无关的协整向量的个数。

如何判断一组  $I(1)$  变量间是否存在协整关系呢？

首先，这些变量必须在理论上可能存在长期均衡关系；否则，进行协整分析将没有意义。

其次，如果只有两个变量，则可以直接画图，看二者的时间趋势是否有相似性。

但此法不严格，也不适用于两个以上的变量。

Engle and Granger (1987) 提出如下 EG-ADF 检验。

不失一般性，考虑两个单位根变量  $\{y_t, x_t\}$  (此检验也适用于多个单位根变量)。

原假设为 $\{y_t, x_t\}$ 存在协整关系，且协整系数为 $\{1, -\theta\}$ 。

如果原假设成立，则 $\{z_t \equiv y_t - \theta x_t\}$ 为平稳过程。

如果 $\theta$ 已知(比如，通过经济理论而知)，则可用 ADF 检验来确定 $\{z_t\}$ 是否平稳。

如果接受“ $\{z_t\}$ 为平稳”，则认为 $\{y_t, x_t\}$ 存在协整关系。

但通常我们不知道 $\theta$ ，故 EG-ADF 检验分两步进行。

**第一步** 用 OLS 估计协整系数 $\theta$ ，即

$$y_t = \phi + \theta x_t + z_t \quad (14.37)$$

在“ $\{y_t, x_t\}$ 存在协整关系”的原假设下，虽然 $\{y_t, x_t\}$ 为非平稳的 I(1)过程，但 $\{z_t\}$ 为平稳过程，故 OLS 的估计量 $\hat{\phi}$ 与 $\hat{\theta}$ 是一致估计量。

**第二步** 对残差序列 $\{\hat{z}_t \equiv y_t - \hat{\phi} - \hat{\theta}x_t\}$ 进行 ADF 检验，确定其是否平稳。

如果检验结果确认 $\{\hat{z}_t\}$ 为平稳，则接受“ $\{y_t, x_t\}$ 存在协整关系”的原假设。估计出的协整关系“ $y_t = \hat{\phi} + \hat{\theta}x_t$ ”即为 $\{y_t, x_t\}$ 之间的长期均衡关系。

反之，如果检验结果认为 $\{\hat{z}_t\}$ 不平稳，则拒绝原假设，认为不存在协整关系。

由于协整系数 $\hat{\theta}$ 是估计出来的，不一定是真实的协整系数，故 EG-ADF 统计量的临界值与普通的 ADF 检验不同，参见表 14.2。

表 14.2 EG-ADF 检验的临界值

方程(14.37)中 解释变量 $x_t$ 的个数	10%	5%	1%
1	-3.12	-3.41	-3.96
2	-3.52	-3.80	-4.36
3	-3.84	-4.16	-4.73
4	-4.20	-4.49	-5.07

资料来源：Stock and Watson (2012, p. 696)。

表 14.2 中的临界值取决于方程(14.37)中解释变量 $x_t$ 的个数。



如果考察两个变量之间的协整关系，则  $x_t$  的个数为 1。

如果考察三个变量之间的协整关系，则  $x_t$  的个数为 2；以此类推。

比如，当只有 1 个  $x_t$  时，显著性为 10% 的临界值为 -3.12，5% 的临界值为 -3.41，而 1% 的临界值为 -3.96。

然后进行左边单侧检验，即拒绝域在临界值的左边。

无论是否有漂移项(常数项)，表 14.2 中的临界值都适用。

EG-ADF 法的缺点是，不能处理同时存在多个协整关系(即协整秩大于 1)的情形。

由于 EG-ADF 法分两步进行，第一步估计的误差被带到第二步中，故不是最有效率的方法。

比 EG-ADF 法更有效率的方法是 MLE，同时估计所有参数。

## 14.8 协整的最大似然估计

假设  $\{y_{1t}, \dots, y_{nt}\}$  都是单位根变量，其中变量个数  $n \geq 2$ 。

记随机向量  $\mathbf{y}_t = (y_{1t} \cdots y_{nt})'$ ，则  $\mathbf{y}_t$  的每个分量都为单位根过程。

考虑  $\mathbf{y}_t$  的向量自回归模型：

$$\mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\Phi}_2 \mathbf{y}_{t-2} + \cdots + \boldsymbol{\Phi}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t \quad (14.38)$$

根据与 ADF 检验类似的推导, 可得此 VAR 模型对应的向量误差修正模型(Vector Error Correction Model, 简记 VECM):

$$\Delta \mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\Gamma}_0 \mathbf{y}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \mathbf{y}_{t-1} + \cdots + \boldsymbol{\Gamma}_{p-1} \Delta \mathbf{y}_{t-p+1} + \boldsymbol{\varepsilon}_t \quad (14.39)$$

其中,  $\boldsymbol{\Gamma}_0 \mathbf{y}_{t-1}$  为误差修正项。当  $\mathbf{y}_{t-1}$  偏离长期均衡关系时,  $\boldsymbol{\Gamma}_0 \mathbf{y}_{t-1}$  即为  $\Delta \mathbf{y}_t$  将要作出的调整(即误差修正)。

矩阵  $\boldsymbol{\Gamma}_0$  的秩即为协整秩。

如果  $\text{rank}(\boldsymbol{\Gamma}_0) = 0$ , 则不存在协整关系。

如果  $\text{rank}(\boldsymbol{\Gamma}_1) = 1$ , 则存在一个协整关系; 以此类推。

假设协整秩为 $h$ ，且扰动项服从 $n$ 维正态分布。

Johansen (1988)使用 MLE 来估计此 VECM 模型。

在满足“ $\text{rank}(\Gamma_0) = h$ ”的条件下，最大化样本数据 $\{\mathbf{y}_1, \dots, \mathbf{y}_T\}$ 的对数似然函数。

为求解此约束极值问题，须先确定协整秩 $h$ 。首先进行以下检验：

$$H_0 : \text{rank}(\Gamma_0) = 0 \quad \text{vs} \quad H_1 : \text{rank}(\Gamma_0) > 0 \quad (14.40)$$

其中，原假设为协整秩为 0，即不存在协整关系；而替代假设为协整秩大于 0，即存在协整关系。

当协整秩为 $h$ 时，系数矩阵 $\boldsymbol{\Gamma}_0$ 有 $h$ 个自由(线性无关)的行向量。

协整秩 $h$ 越大，则对矩阵 $\boldsymbol{\Gamma}_0$ 的约束越少，其对应的似然函数最大值应该越大。

据此可进行似然比检验，由于此检验统计量涉及矩阵的迹(主对角线元素之和)，故称为迹统计量(trace statistics)，记为 $\lambda_{\text{trace}}$ 。

由于迹检验(trace test)是似然比检验，故为单边右侧检验，即 $\lambda_{\text{trace}}$ 越大，则越倾向于拒绝原假设。

如果接受“ $H_0 : \text{rank}(\boldsymbol{\Gamma}_0) = 0$ ”，则认为不存在协整关系。

反之，则继续检验是否存在多个协整关系：

$$H_0 : \text{rank}(\boldsymbol{\Gamma}_0) = 1 \quad \text{vs} \quad H_1 : \text{rank}(\boldsymbol{\Gamma}_0) > 1 \quad (14.41)$$

其中，原假设为协整秩为 1，即仅存在一个协整关系；而替代假设为协整秩大于 1，即存在多个协整关系。

依此顺序不断进行检验，直到接受  $H_0$ ，确认协整秩  $h$  为止。

Johansen (1988) 还考虑了另一类检验：

$$H_0 : \text{rank}(\boldsymbol{\Gamma}_0) = h \quad \text{vs} \quad H_1 : \text{rank}(\boldsymbol{\Gamma}_0) = h + 1 \quad (14.42)$$

其中，原假设为协整秩为  $h$ ，即存在  $h$  个协整关系；而替代假设为协整秩等于  $(h + 1)$ ，即存在  $(h + 1)$  个协整关系。

此检验的统计量为最大特征值统计量(maximum eigenvalue statistics), 记为 $\lambda_{\max}$ , 称为最大特征值检验(maximum eigenvalue test)。

一般认为, 迹检验的效果比特征值检验更好, 故前者为 Stata 的默认方法。

确认协整秩 $h$ 后, 即可在给定“ $\text{rank}(\Gamma_0) = h$ ”的条件下, 使用 MLE 估计 VECM 模型。如果存在时间趋势与季节效应(为季度数据或月度数据), 则 VECM 模型变为:

$$\Delta \mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\Gamma}_0 \mathbf{y}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \mathbf{y}_{t-1} + \cdots + \boldsymbol{\Gamma}_{p-1} \Delta \mathbf{y}_{t-p+1} + \boldsymbol{\delta} t + \mathbf{w}_1 s_1 + \cdots + \mathbf{w}_m s_m + \boldsymbol{\varepsilon}_t$$

(14.43)

其中， $\delta t$  为时间趋势项，而  $\{s_1, \dots, s_m\}$  为季节虚拟变量。

仍可进行条件 MLE 估计。

## 14.9 协整分析的 Stata 实例

协整分析的起点是所有相关变量都是单位根变量，为此首先须进行 ADF 检验。

如果确定所有变量都为单位根过程，则可进一步检验这些  $I(1)$  变量是否存在协整关系。



## (1) 检验协整秩

检验协整秩的 Stata 命令为(假设变量为  $x, y, z$ )

```
. vecrank x y z, lags(#) max trend(none) trend(trend)
```

选择项“lags(#)”表示对应的 VAR 模型(14.38)中滞后的阶数, 默认为“lags(2)”。

选择项“max”表示也进行最大特征值检验, 默认仅进行迹检验。

选择项“trend(none)”表示不包括常数项或时间趋势。

选择项 “trend(trend)” 表示包括常数项与时间趋势；默认包括常数项，但不包括时间趋势。

命令 vecrank 的输出结果将列出 “ $h = 0, 1, \dots, n-1$ ” 的一系列检验，并以星号(\*)标出所接受 $h$ 值，即协整秩。

## (2) 估计协整关系

在作完协整秩检验后，如果确定 $h = 0$ ，则认为这些变量不存在协整关系。

反之，如果确定 $h \geq 1$ ，则可对 VECM 模型进行最大似然估计(也可使用 Engle-Granger 两步法，但效率更低，且在 $h \geq 2$ 的情况下无法使用)。

使用 MLE 估计 VECM 模型的 Stata 命令为

```
.   vec   x   y   z, lags(#)   rank(#)   trend(none)  
trend(trend) sindicators(varlist)
```

选择项“lags(#)”表示对应的 VAR 模型(14.38)中滞后的阶数，默认为“lags(2)”。

选择项“rank(#)”表示协整秩的阶数，默认为“rank(1)”。  
选择项“trend(none)”表示不包括常数项或时间趋势。

选择项“trend(trend)”表示包括常数项与时间趋势；默认包括常数项，但不包括时间趋势。

选择项“sindicators(varlist)”表示加入季节虚拟变量。

### (3) 诊断性检验

估计完 VECM 模型后，应对模型的假设进行诊断性检验 (diagnostic checking)，主要包括残差有无自相关，以及模型的平稳性。

```
. vec1mar
```

此命令对残差是否存在自相关进行  $LM$  检验。

```
. vecstable, graph
```

此命令检验 VECM 系统是否为平稳过程。如果所有特征值都在单位圆内部，则为平稳过程。

选择项 “graph” 表示画出特征值的几何分布图。

#### (4) 脉冲响应函数与预测

估计完 VECM 模型后，也可计算脉冲响应函数或进行预测，其命令与 VAR 模型相同。

对于协整分析而言，一般主要关注长期的均衡关系(即协整关系)，而不太关心短期的调整过程。

以数据集 `mpyr.dta` 为例，对美国的货币需求函数进行协整分析。

该数据集包含了美国 1900—1989 年的以下年度宏观变量：  
 $\log p$ (价格水平的对数)， $\log y$ (名义净国民生产总值的对数)， $\log m1$ ( $M_1$  的对数)， $\log mr$ (实际货币的对数，即  $\log m1 - \log p$ )，以及  $r$ (名义利率)。

从经济理论出发, 可将货币需求函数写为

$$\log m r_t = \beta_0 + \beta_1 \log y_t + \beta_2 r_t + \varepsilon_t \quad (14.44)$$

其中,  $\beta_1$  为货币需求的收入弹性, 一般认为接近于 1

$\beta_2$  为货币需求的利率“半弹性”(semielasticity), 一般为负。

假定以上变量均为单整 I(1) 过程(参见习题), 故应进行协整分析。

首先, 从图形上大致考察  $(\log m r, \log y, r)$  是否存在协整关系(参见图 14.6):

```
. use mpyr.dta,clear  
  
. tsline logmr logy r,lp(solid dash shortdash)  
xlabel(1900(10)1990)
```

选择项 “`lp(solid dash shortdash)`” 表示分别用实线、虚线与短虚线来画图。

选择项 “`xlabel(1900(10)1990)`” 表示，在横轴从 1900-1990 年，每隔 10 年做个标注(label)。

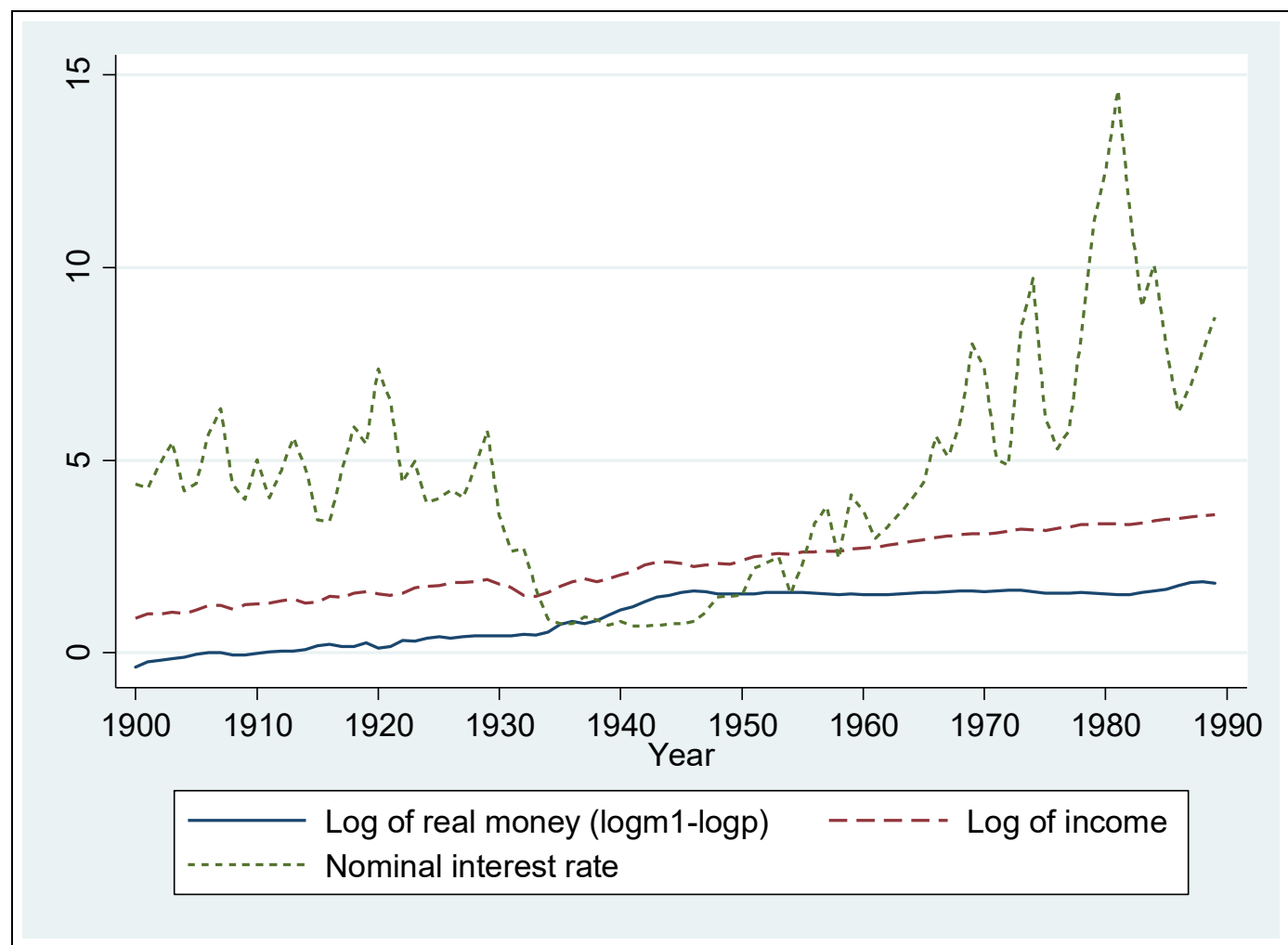


图 14.6 实际货币与收入的时间趋势



实际货币对数与收入对数的时间走势比较接近，而名义利率似乎与实际货币对数呈反向变动。

$(\log mr, \log y, r)$ 可能存在长期均衡关系，即为协整系统。

这三个变量似乎都存在时间趋势。

首先，需要确定此系统的协整秩，即究竟有多少个线性无关的协整关系。

对于这个三个变量的系统，可能的协整秩为 0，1 或 2。

在使用命令 `vecrank` 进行协整秩检验时，需要指定相应 VAR 模型的滞后阶数。

先检验该系统所对应的 VAR 模型的滞后阶数：

```
. varsoc logmr logy r
```

Lag-order selection criteria								
Sample: 1904 thru 1989						Number of obs = 86		
Lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	-251.056				.073876	5.90827	5.94272	5.99388
1	132.578	767.27	9	0.000	.000012	-2.80415	-2.66632	-2.46168*
2	148.293	31.429	9	0.000	.00001*	-2.96029*	-2.7191*	-2.36098
3	151.979	7.3723	9	0.598	.000012	-2.83672	-2.49215	-1.98055
4	162.506	21.054*	9	0.012	.000011	-2.87222	-2.42429	-1.75921

\* optimal lag  
 Endogenous: logmr logy r  
 Exogenous: \_cons

上表中打星号者为根据不同准则所选择的滞后阶数。其中，AIC 准则选择滞后二阶，而 BIC 准则选择滞前一阶。

为了保守起见，在此选择滞后二阶。

下面进行协整秩检验：

```
. vecrank logmr logy r,lags(2) trend(trend) max
```

选择项“lags(2)”表示对应的 VAR 模型滞后二阶(也是默认值)。

选择项“trend(trend)”表示既包括常数项，也包括时间趋势项。

选择项“max”表示显示最大特征值统计量。

Johansen tests for cointegration					
Trend: Linear			Number of obs = 88		
Sample: 1902 thru 1989			Number of lags = 2		
<hr/>					
Maximum				Trace	Critical
rank	Params	LL	Eigenvalue	statistic	value
0	15	138.03791	.	46.3731	5%
1	20	153.13651	0.29047	16.1759*	18.17
2	23	160.58579	0.15575	1.2773	3.74
3	24	161.22445	0.01441		
<hr/>					
Maximum			———Eigenvalue———		Critical
rank	Params	LL		Maximum	value
0	15	138.03791	.	30.1972	5%
1	20	153.13651	0.29047	14.8985	16.87
2	23	160.58579	0.15575	1.2773	3.74
3	24	161.22445	0.01441		
<hr/>					
* selected rank					

迹检验(trace statistic)结果表明，只有一个线性无关的协整向量(上表中打星号“\*”者)。

最大特征值检验(max statistic)表明，可以在 5%的水平上拒绝“协整秩为 0”的原假设( $30.1972 > 23.78$ )，但无法拒绝“协整秩为 1”的原假设( $14.8985 > 16.87$ )。

因此，选择协整秩为 1。

使用 Johansen 的 MLE 方法估计该系统的向量误差修正模型 (VECM):

```
.vec logmr logy r,lags(2) rank(1)
```

选择项 “lags(2)” 表示对应的 VAR 模型滞后二阶(也是默认值)

选择项 “rank(1)” 表示协整秩为 1(也是默认值)。

Vector error-correction model						
Sample: 1902 thru 1989			Number of obs		=	88
			AIC		=	-3.037506
Log likelihood = 150.6503			HQIC		=	-2.8447
Det(Sigma_ml) = 6.54e-06			SBIC		=	-2.55893
Equation	Parms	RMSE	R-sq	chi2	P>chi2	
D_logmr	5	.050841	0.2758	31.61646	0.0000	
D_logy	5	.056773	0.3483	44.35204	0.0000	
D_r	5	1.1287	0.2133	22.50845	0.0004	
	Coefficient	Std. err.	z	P> z	[95% conf. interval]	
D_logmr						
_cel						
L1.	-.0533601	.0410678	-1.30	0.194	-.1338515	.0271314
logmr						
LD.	.2079032	.1107256	1.88	0.060	-.0091149	.4249214
logy						
LD.	.0086587	.101984	0.08	0.932	-.1912263	.2085438
r						
LD.	-.0063968	.0052854	-1.21	0.226	-.0167559	.0039624
_cons	.0186333	.0064766	2.88	0.004	.0059393	.0313273

D_logy							
	_cel						
	L1.	.0298268	.0458591	0.65	0.515	-.0600554	.119709
	logmr						
	LD.	.2666361	.1236437	2.16	0.031	.0242989	.5089733
	logy						
	LD.	.2330244	.1138823	2.05	0.041	.0098191	.4562296
	r						
	LD.	-.0145323	.005902	-2.46	0.014	-.0261001	-.0029646
	_cons	.0157173	.0072323	2.17	0.030	.0015424	.0298923
D_r							
	_cel						
	L1.	-3.482578	.9117297	-3.82	0.000	-5.269536	-1.695621
	logmr						
	LD.	2.663613	2.458173	1.08	0.279	-2.154318	7.481544
	logy						
	LD.	.6533844	2.264106	0.29	0.773	-3.784182	5.09095
	r						
	LD.	.294868	.1173386	2.51	0.012	.0648885	.5248475
	_cons	-.0001509	.1437852	-0.00	0.999	-.2819648	.281663



Cointegrating equations

Equation	Parms	chi2	P>chi2
_cel	2	794.1155	0.0000

Identification: beta is exactly identified

Johansen normalization restriction imposed

beta	Coefficient	Std. err.	z	P> z	[95% conf. interval]		
_cel	logmr	1	.	.	.	.	
	logy	-.9754246	.0346169	-28.18	0.000	-1.043273	-.9075767
	r	.1124051	.0097191	11.57	0.000	.093356	.1314542
	_cons	.7299535	.	.	.	.	.

此表上部为误差修正模型(Vector error-correction model)，下部为协整方程(Cointegrating equation)，以“\_cel”来表示。

我们主要对货币需求函数感兴趣，即协整方程所代表的长期均衡关系。

根据此协整方程，协整向量为(1, -0.98, 0.11)，其中  $\log m_r$  的系数被标准化为 1，故其标准误为缺失；而其他两个变量( $\log y$  与  $r$ )的协整系数均在 1%水平上显著。

将协整向量移项，可将估计的货币需求函数写为

$$\widehat{\log m r_t} = -0.73 + 0.98 \log y_t - 0.11 r_t \quad (14.45)$$

其中，货币需求的收入弹性为 0.98，而货币需求的利率半弹性为-0.11，均符合经济理论的预期。

下面检验 VECM 模型的残差是否存在自相关。如果存在自相关，则需增加滞后阶数。

```
. vec1mar
```

Lagrange-multiplier test			
lag	chi2	df	Prob > chi2
1	6.6260	9	0.67599
2	12.5541	9	0.18384

H0: no autocorrelation at lag order

可以接受“无自相关”的原假设。

下面检验此 VECM 系统是否稳定，结果如图 14.7:

. vecstable,graph

Eigenvalue stability condition

Eigenvalue	Modulus
1	1
1	1
.4092107 + .4061819 <i>i</i>	.576574
.4092107 - .4061819 <i>i</i>	.576574
.2217304 + .07266624 <i>i</i>	.233334
.2217304 - .07266624 <i>i</i>	.233334

The VECM specification imposes 2 unit moduli.

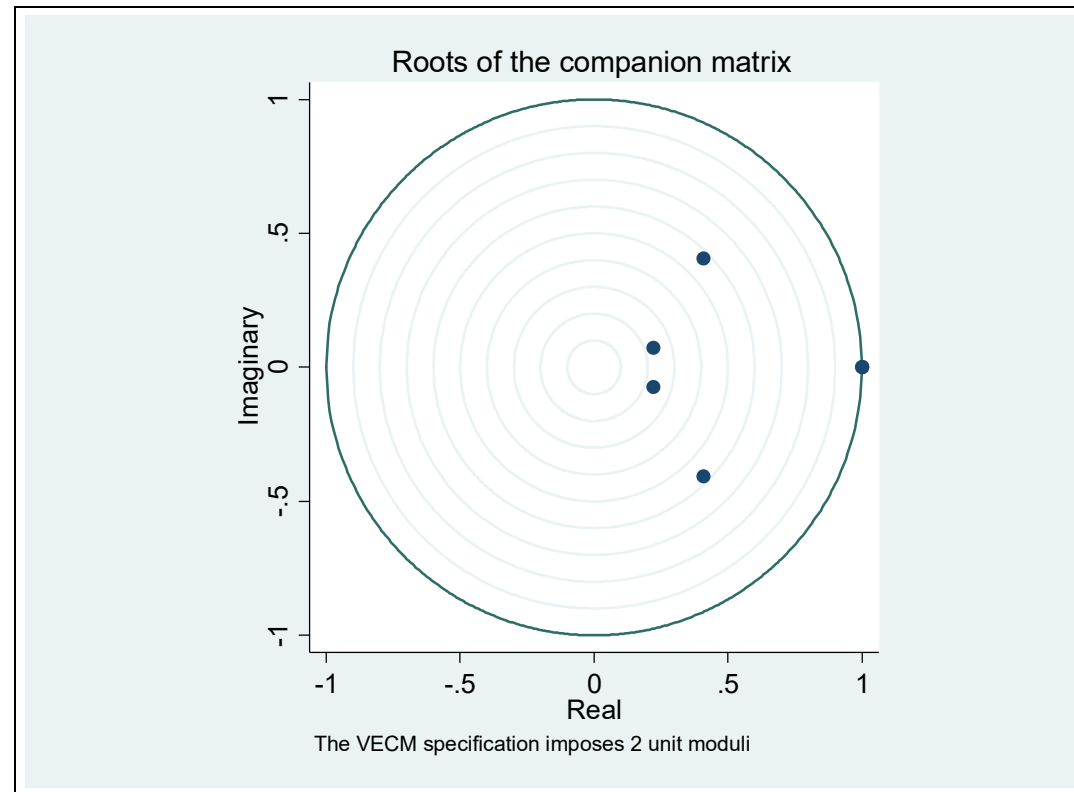


图 14.7 VECM 系统稳定性的判别图

除了 VECM 模型本身所假设的单位根之外(The VECM specification imposes 2 unit moduli), 伴随矩阵的所有特征值均落在单位圆之内, 故是稳定的系统。

作为对比，下面用 OLS 估计此长期均衡关系(EG-ADF 两步法):

```
. reg logmr logy r
```

Source	SS	df	MS	Number of obs	=	90
Model	41.216226	2	20.608113	F(2, 87)	=	1169.93
Residual	1.53248421	87	.017614761	Prob > F	=	0.0000
				R-squared	=	0.9642
				Adj R-squared	=	0.9633
Total	42.7487102	89	.480322587	Root MSE	=	.13272
logmr	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
logy	.9418376	.0196178	48.01	0.000	.9028451	.9808302
r	-.0832229	.0053829	-15.46	0.000	-.0939219	-.0725239
_cons	-.7737089	.0426886	-18.12	0.000	-.8585572	-.6888606

OLS 系数估计值与 Johansen 的 MLE 估计结果比较接近。从理论来说，MLE 估计更有效率。