

第 13 章 平稳时间序列

“时间序列数据”分为“平稳序列”(stationary)与“非平稳序列”(non-stationary)两大类，需使用不同的计量方法。

本章介绍平稳序列，下一章介绍非平稳序列。

13.1 时间序列的自相关

时间序列指同一个体在不同时点上的观测数据。比如，在 1978-2013 年期间，中国每年的国内生产总值。

对于离散时间 $\{1, 2, \dots, T\}$ ，可将时间序列写为 $\{y_1, y_2, \dots, y_T\}$ ，其中每个 y_t 都是随机变量。

时间序列的最大特点是存在自相关，不同期的观测值之间存在相关性。

定义 时间序列 $\{y_t\}$ 的 k 阶自协方差(autocovariance of order k)为

$$\gamma_k \equiv \text{Cov}(y_t, y_{t+k}) = E[(y_t - \mu)(y_{t+k} - \mu)] \quad (13.1)$$

其中， $\mu \equiv E(y)$ 为总体均值。

γ_k 反映同一变量(y)相隔 k 期之间的自相关程度。

当 $k = 0$ 时， $\gamma_0 \equiv \text{Var}(y)$ 。

对 γ_k 的估计值为样本自协方差:

$$\hat{\gamma}_k \equiv \frac{1}{T-k} \sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y}) \quad (13.2)$$

其中, $\bar{y} \equiv \frac{1}{T} \sum_{t=1}^T y_t$ 为样本均值。

自协方差受变量单位的影响。为此, 将其标准化。

定义 时间序列 $\{y_t\}$ 的 k 阶自相关系数(autocorrelation of order k) 为

$$\rho_k \equiv \text{Corr}(y_t, y_{t+k}) \equiv \frac{\text{Cov}(y_t, y_{t+k})}{\text{Var}(y_t)} \quad (13.3)$$

自相关系数 ρ_k 将自协方差 γ_k 标准化为介于 $[-1, 1]$ 之间的量。

对于严格平稳过程， ρ_k 不依赖于具体时间，仅是滞后阶数 k 的函数，称为“自相关函数”(Autocorrelation Function，简记 ACF)。

将 (k, ρ_k) 画成图，即为“自相关图”(correlogram)。

对 ρ_k 的估计值为样本自相关系数：

$$\hat{\rho}_k \equiv \hat{\gamma}_k / \hat{\gamma}_0 \quad (13.4)$$

其中， $\hat{\gamma}_0 \equiv \frac{1}{T-1} \sum_{t=1}^T (y_t - \bar{y})^2$ 为样本方差。

这些数字特征是时间序列固有的特征，不依赖于模型设定。

在设定模型时，应尽可能与这些数字特征一致。

例 使用数据集 `gdp_china.dta` 考察 1978-2013 年, 中国国内生产总值(1978 年不变价格, 亿元), 记为 y 。

定义时间变量后, 看 GDP 的时间趋势(参见图 13.1)。

```
. use gdp_china.dta, clear  
  
. tsset year  
  
. tsline y, xlabel(1980(10)2010)
```

其中, “`tsline`” 表示画时间趋势图, 在此等价于命令 “`line gdp year`” (`year` 为时间变量)。“`xlabel(1980(10)2010)`” 表示在横轴 1980-2010 期间, 每隔 10 年做个标注(label)。

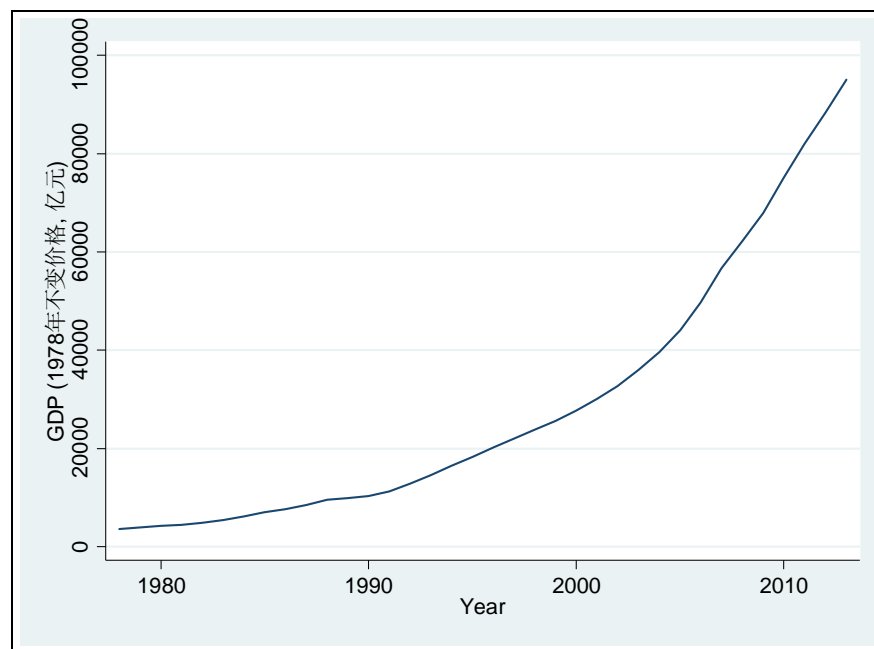


图 13.1 GDP 的指数增长趋势(1978-2013)

GDP 存在指数增长(exponential growth)的趋势。通常的处理方法是，将 GDP 取对数，把指数趋势变为线性趋势。

计算 GDP 对数，再次画时间趋势图(参见图 13.2)。

```
. gen lny=log(y)  
. tsline lny,xlabel(1980(10)2010)
```

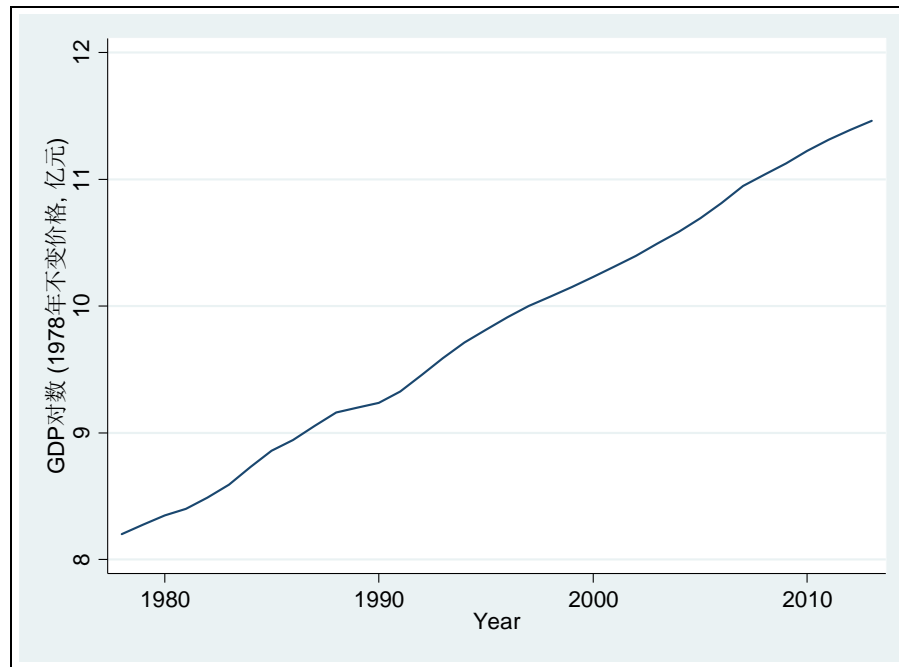


图 13.2 GDP 对数的线性增长趋势(1978-2013)

GDP 对数存在线性趋势，但依然不平稳(期望值不断增长)。

将 GDP 对数进行一阶差分，然后画时间趋势图。

```
. gen dlnty=d.lnty  
. tsline dlnty,xlabel(1980(10)2010)
```

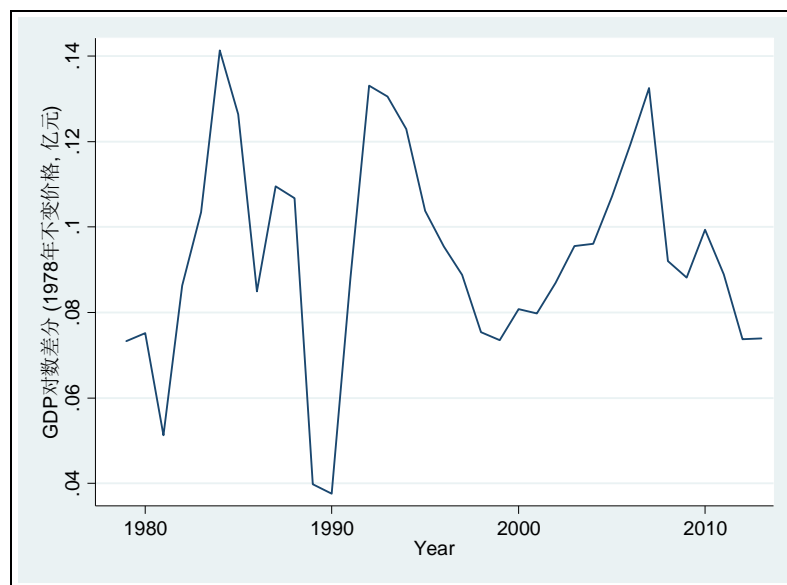


图 13.3 GDP 对数差分的时间趋势(1978-2013)

$\Delta \ln y_t$ 不存在明显的时间趋势，可大致视为平稳序列。

之所以考察 GDP 对数差分，因为它约等于 GDP 的增长率：

$$\begin{aligned}\Delta \ln y_t &\equiv \ln y_t - \ln y_{t-1} = \ln \left(\frac{y_t}{y_{t-1}} \right) \\ &= \ln \left(\frac{y_{t-1} + \Delta y_t}{y_{t-1}} \right) = \ln \left(1 + \frac{\Delta y_t}{y_{t-1}} \right) \approx \frac{\Delta y_t}{y_{t-1}}\end{aligned}\tag{13.5}$$

其中，根据泰勒展开的一阶近似，当 $x \approx 0$ 时， $\ln(1+x) \approx x$ 。

有时直接将 $\Delta \ln y_t$ 视为 y_t 的增长率。如增长率较高，则误差较大。

直接计算 GDP 的增长率(记为 g), 并与 GDP 对数差分进行画图比较(参见图 13.4)。

```
. gen g=(y-l.y)/l.y  
(1 missing value generated)  
. tsline dlny g,xlabel(1980(10)2010)  
lpattern(dash)
```

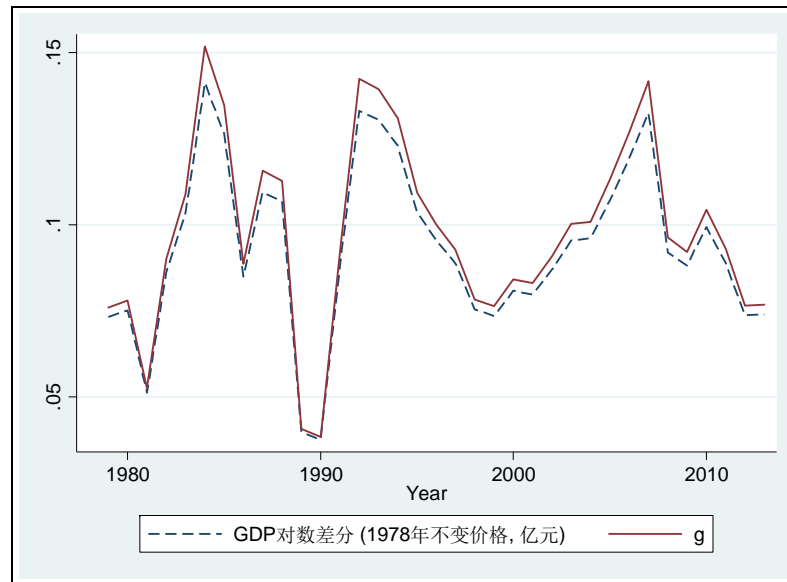
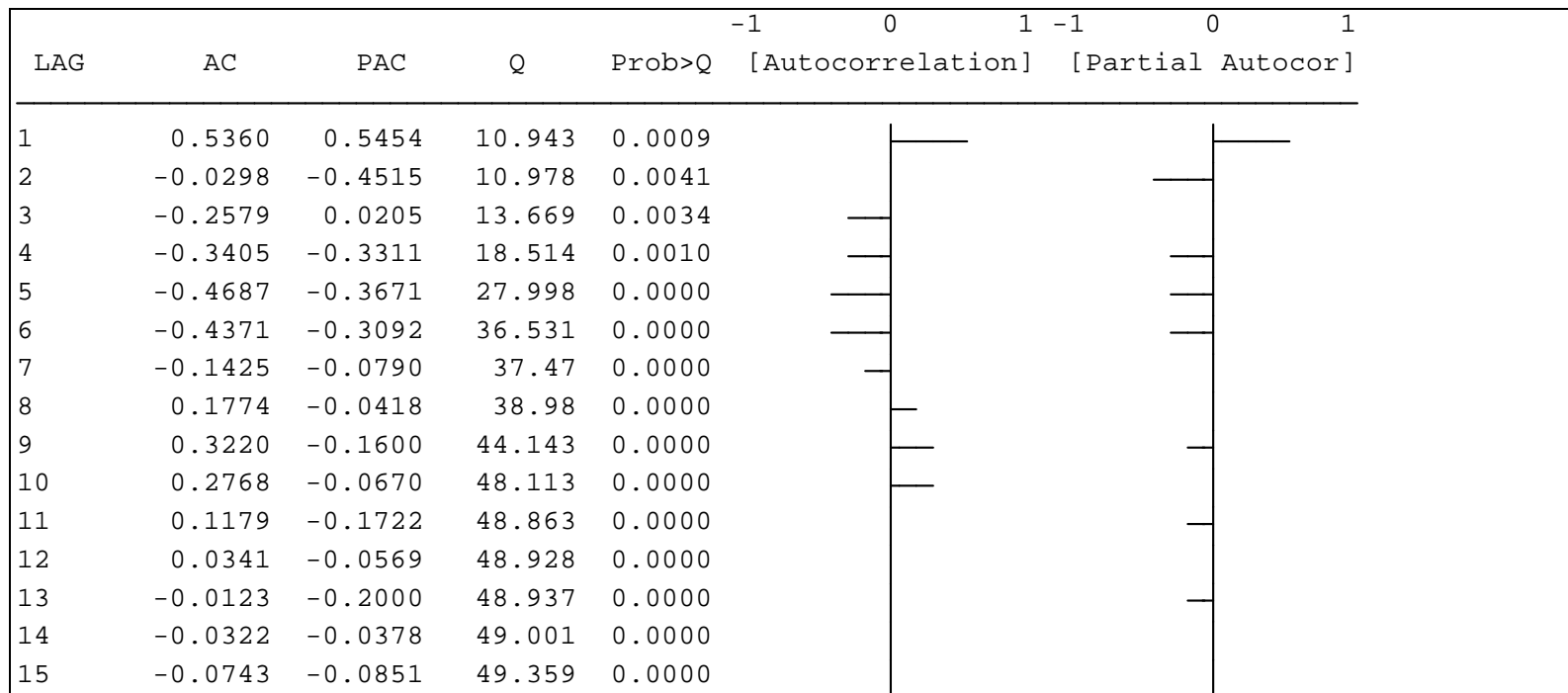


图 13.4 GDP 增长率的两种计算方法(1978-2013)

通过自相关图，考察 GDP 对数差分的各阶自相关系数。

. corrgram dlny

其中，“corrgram”表示 correlogram，即画自相关图。



使用画自相关图的另一命令。

```
. ac dlny, lags(20)
```

其中，“ac”表示 autocorrelation ；

选择项 “lags(20)” 表示画 1-20 阶的自相关图；

默认所画的最高阶数为 $\min\{\text{floor}(n/2) - 2, 40\}$ ，其中 $\text{floor}(n/2)$ 为不超过 $n/2$ 的最大整数。参见图 13.5。

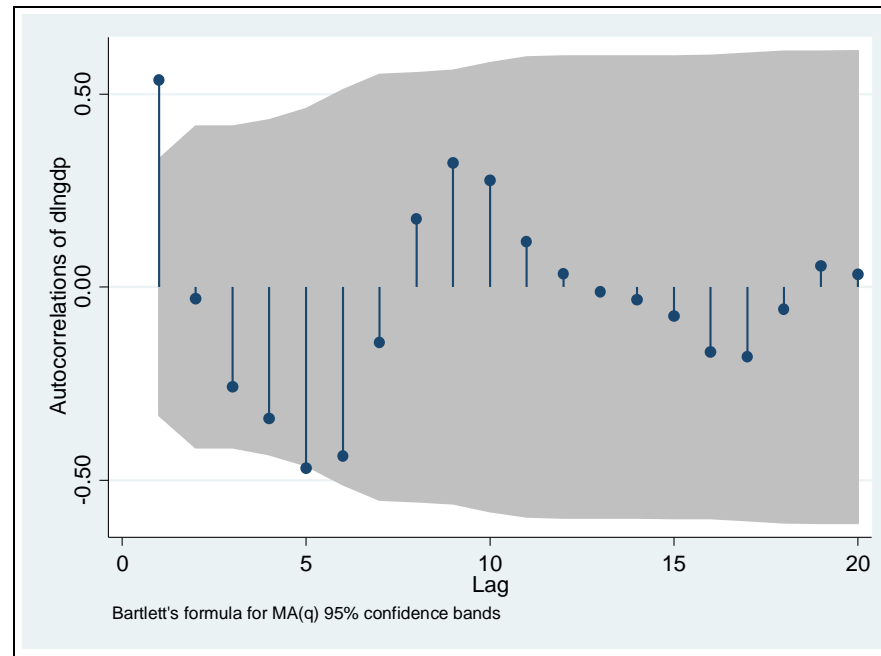


图 13.5 GDP 对数差分的自相关图

一阶与五阶自相关系数显著不为 0，其他阶不显著。

13.2 一阶自回归

此前均强调以回归模型推断因果关系。

从客户角度仅关心某变量(比如股价)的未来值,可用该变量的过去值来预测其未来值(因为时间序列一般存在自相关)。

这种模型称为“单变量时间序列”(univariate time series)。

此时可不必理会因果关系,只考虑相关关系即可。

比如,看到街上有人带伞,可预测今天下雨,但行人带伞并不导致下雨。

最简单的预测方法为，使用过去值预测当前值，即一阶自回归模型(AR(1)):

$$y_t = \beta_0 + \beta_1 y_{t-1} + \varepsilon_t \quad (t = 2, \dots, T) \quad (13.6)$$

其中，扰动项 ε_t 为白噪声，故 $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0, \quad \forall t \neq s$ 。

假设自回归系数 $|\beta_1| < 1$ ，则 $\{y_t\}$ 为渐近独立的平稳过程。

由于 y_{t-1} 依赖于 $\{\varepsilon_{t-1}, \dots, \varepsilon_1\}$ ，而扰动项 ε_t 与 $\{\varepsilon_{t-1}, \dots, \varepsilon_1\}$ 不相关，故 y_{t-1} 为前定变量，与 ε_t 不相关，故 OLS 一致。

使用 OLS 将损失一个样本容量。

为提高估计效率，可使用 MLE (须假设扰动项服从正态分布)。

继续上例，以 OLS 估计 $\Delta \ln y_t$ 的一阶自回归模型。

仅使用 2013 年前的数据回归，然后预测 2013 年的 GDP。

```
. reg dlny l.dlny if year<2013,r
```

由于假设扰动项 ε_t 无自相关，故使用异方差稳健的标准误即可，不必使用异方差自相关稳健的 HAC 标准误。

Linear regression					Number of obs = 33	
					F(1, 31) = 12.99	
					Prob > F = 0.0011	
					R-squared = 0.2879	
					Root MSE = .02147	
dlny	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
dlny L1.	.5362727	.1487888	3.60	0.001	.2328159	.8397295
_cons	.0437698	.0144049	3.04	0.005	.0143908	.0731488

可得如下回归方程(常数项与斜率均在 1% 水平上显著):

$$\widehat{\Delta \ln y_t} = 0.0437698 + 0.5362727 \Delta \ln y_{t-1} \quad (13.7)$$

计算回归方程的拟合值, 即 $\widehat{\Delta \ln y_t}$, 并记为 dlny1。

```
. predict dlny1
(option xb assumed; fitted values)
(2 missing values generated)
```

```
. list dlny1 if year==2013
```

	dlny1
36.	.083309

因此， $\widehat{\Delta \ln y_{2013}} = 0.083309$ 。

由于 $\widehat{\ln y_{2013}} = \ln y_{2012} + \widehat{\Delta \ln y_{2013}}$ ，故 2013 年 GDP 的预测值为

$$\widehat{y_{2013}} = \exp\left(\ln y_{2012} + \widehat{\Delta \ln y_{2013}}\right)。$$

在 Stata 中, 使用 “ $x[n]$ ” 表示变量 x 的第 n 个观测值, 故可计算如下:

```
. dis exp(lny[35]+dlny1[36])
```

```
95985.114
```

其中, “ $\ln y[35]$ ” 表示变量 $\ln y$ 的第 35 个观测值(即 2012 年), 而 “ $d\ln y1[36]$ ” 表示变量 $d\ln y1$ 的第 36 个观测值(即 2013 年), 因为样本容量为 36。

根据 AR(1)模型, 2013 年 GDP 的预测值为 95,985.114 亿元(1978 年不变价格)。

对比 2013 年的实际 GDP, 并计算预测误差, 即 $(y_{2013} - \widehat{y_{2013}})$:

```
. dis y[36]
```

```
95089.211
```

```
. dis y[36]-exp(lny[35]+dlny1[36])
```

```
-895.90347
```

预测误差为-895.90347 亿元，高估了 895.90347 亿元。

13.3 高阶自回归

在 AR(1)模型中，假设扰动项无自相关，故 OLS 一致。

如模型为 AR(2)，被误设为 AR(1)，则二阶滞后项 $\beta_2 y_{t-2}$ 被纳入扰动项：

$$y_t = \beta_0 + \beta_1 y_{t-1} + (\beta_2 y_{t-2} + \varepsilon_t) \quad (13.8)$$

由于扰动项为 $(\beta_2 y_{t-2} + \varepsilon_t)$ ，故与 y_{t-1} 相关，OLS 不一致；须引入 $\beta_2 y_{t-2}$ 才能得到一致估计。

从预测的角度，高阶滞后项可能包含有用信息。

考虑 p 阶自回归模型，记为 AR(p):

$$y_t = \beta_0 + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + \varepsilon_t \quad (13.9)$$

其中，扰动项 ε_t 为白噪声(无自相关)，故 OLS 一致。

通常不知道滞后期 p 。如何估计 \hat{p} ？

方法一：

设最大滞后期 p_{\max} ，令 $\hat{p} = p_{\max}$ 进行估计，对最后一个滞后期系数的显著性进行 t 检验。

如接受该系数为 0，令 $\hat{p} = p_{\max} - 1$ ，重新估计，再对(新的)最后一个滞后期的系数进行 t 检验，如显著，则停止；否则，令 $\hat{p} = p_{\max} - 2$ ；以此类推。

此准则称为“由大到小的序贯 t 规则” (general-to-specific sequential t rule)。

方法二：

使用信息准则，选择 \hat{p} 使 AIC 或 BIC 最小化，分别记为 \hat{p}_{AIC} 与 \hat{p}_{BIC} 。比如，

$$\min_p \text{AIC} \equiv \ln\left(\frac{\text{SSR}}{T}\right) + \frac{2}{T}(p+1) \quad (13.10)$$

其中，SSR 为残差平方和。

\hat{p}_{BIC} 是真实滞后阶数 p 的一致估计， \hat{p}_{AIC} 在大样本中可能高估 p 。

在小样本中，这两种信息准则难分优劣，都很常用。

实践中，可结合以上两种方法来确定 \hat{p} 。

如二者结果不一致,为了保守起见(尽量避免遗漏变量偏差),可取二者滞后阶数的大者。

还可检验模型的残差是否存在自相关(比如,使用 Q 检验);如果残差存在自相关,则须扩大滞后阶数。

回到上节 GDP 对数差分的例子。

首先,使用信息准则确定滞后阶数 p 。

```
. quietly reg dlny l.dlny if year<2013,r  
. estat ic
```


Akaike's information criterion and Bayesian information criterion						
Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	33	75.35938	80.96115	2	-157.9223	-154.9293
Note: N=Obs used in calculating BIC; see <u>[R] BIC note</u>						

AR(1)的 AIC 为-157.9223, BIC 为-154.9293。

估计 AR(2)模型, 并计算信息准则。

```
. reg dlpy l(1/2).dlpy if year<2013,r
```

其中, “l(1/2).dlpy” 表示变量 dlpy 的 1-2 阶滞后。

Linear regression				Number of obs = 32		
				F(2, 29) = 17.51		
				Prob > F = 0.0000		
				R-squared = 0.4234		
				Root MSE = .01979		
dlny	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
dlny						
L1.	.7711595	.1304462	5.91	0.000	.5043671	1.037952
L2.	-.4487175	.1530057	-2.93	0.007	-.7616494	-.1357857
_cons	.0641134	.0128498	4.99	0.000	.0378326	.0903943

dlny 的二阶滞后 L2.dlny 依然在 1% 水平上显著，故根据序贯 t 规则，滞后阶数 p 应至少大于或等于 2。

. estat ic

Akaike's information criterion and Bayesian information criterion						
Model	Obs	ll(null)	ll(model)	df	AIC	BIC
.	32	72.88943	81.69936	3	-157.3987	-153.0015
Note: N=Obs used in calculating BIC; see <u>[R] BIC note</u>						

AR(2)的 AIC 为-157.3987, BIC 为-153.0015; 均比 AR(1)略有上升。故根据信息准则, 应选择 $p=1$, 即 AR(1)模型。

进一步估计 AR(3)模型。

```
. reg dlnt 1(1/3).dlnt if year<2013,r
```

Linear regression					Number of obs = 31	
					F(3, 27) = 12.82	
					Prob > F = 0.0000	
					R-squared = 0.4459	
					Root MSE = .01907	
dlny	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
dlny						
L1.	.7557034	.1261359	5.99	0.000	.4968939	1.014513
L2.	-.4943798	.1785691	-2.77	0.010	-.8607733	-.1279864
L3.	.0204783	.1711994	0.12	0.906	-.3307938	.3717504
_cons	.0692315	.0154137	4.49	0.000	.0376052	.1008579

dlny 的三阶滞后很不显著，根据序贯 t 规则，应选择 $p = 2$ 。

综合以上结果，为避免遗漏变量偏差，应按照序贯 t 规则选择AR(2)模型。

使用命令 `corrgram` 对残差进行 Q 检验也表明, AR(1)的残差存在自相关, 而 AR(2)的残差无自相关(参见习题)。

使用 AR(2)模型预测 GDP, 并与 AR(1)的预测效果对比。

```
. quietly reg dlny l(1/2).dlny if year<2013,r  
  
. predict dlny2  
(option xb assumed; fitted values)  
(3 missing values generated)  
  
. dis exp(lny[35]+dlny2[36])  
95769.998  
  
. dis y[36]-exp(lny[35]+dlny2[36])  
-680.78688
```

对于 2013 年的 GDP, AR(2)模型的预测误差为-680.78688 亿元, 即高估了 680.78688 亿元;

AR(1)模型则高估了 895.90347 亿元。

AR(2)的预测效果优于 AR(1), 因为二阶滞后仍包含有用信息。

13.4 自回归分布滞后模型

在 $AR(p)$ 模型中, 为了提高预测力, 也可引入其他解释变量, 构成“自回归分布滞后模型”(Autoregressive Distributed Lag Model, 简记 $ADL(p, q)$ 或 $ARDL(p, q)$):

$$y_t = \beta_0 + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + \gamma_1 x_{t-1} + \cdots + \gamma_q x_{t-q} + \varepsilon_t \quad (13.11)$$

p 为 y 的自回归阶数, 而 q 为 x 的滞后阶数。

假定扰动项 ε_t 为白噪声, 则 OLS 一致。

对于滞后阶数 (p, q) 的选择, 可使用信息准则(AIC 或 BIC), 或进行序贯检验, 即使用 t 或 F 检验来检验最后一阶系数的显著性。

在 ADL 模型中, 也可引入更多的解释变量; 比如, 变量 z 的 r 阶滞后 $(z_{t-1}, \dots, z_{t-r})$ 。

对于 ADL 模型(13.11), 解释变量 x_{t-1} 对于 y_t 的边际效应为 γ_1 , 但这并非长期效应。

由于 $\{y_t\}$ 与 $\{x_t\}$ 为平稳序列, 故均值不随时间而变, 分别记为 y^* 与 x^* 。

将方程(13.11)两边同时求期望：

$$y^* = \beta_0 + \beta_1 y^* + \cdots + \beta_p y^* + \gamma_1 x^* + \cdots + \gamma_q x^* \quad (13.12)$$

整理可得

$$(1 - \beta_1 - \cdots - \beta_p) y^* = \beta_0 + (\gamma_1 + \cdots + \gamma_q) x^* \quad (13.13)$$

x^* 增加一单位对 y^* 的边际效应为

$$\frac{dy^*}{dx^*} = \frac{\gamma_1 + \cdots + \gamma_q}{1 - \beta_1 - \cdots - \beta_p} \quad (13.14)$$

这就是 x 永久性增加一单位对 y 的长期效应，也称“长期乘数” (long-run multiplier)。

例 Chen(2015)研究中原王朝与北方游牧民族边界纬度(border)的决定因素。

时间序列数据集 border.dta 以每十年作为观测单位(时间变量为decade)，从公元前 221 年秦朝建立至 1911 年清朝灭亡，共有 213 个观测值。

主要解释变量：中原王朝早于游牧政权建立的年数(diff)，中国北方在十年中发生旱灾的年数比例(drought)。

其他控制变量：中原王朝的绝对年龄(age)，游牧对手数目(rival)，中原是否在长城的有效保护之下(wall)，中国是否统一(unified)。

信息准则与序贯规则均支持 ADL(2, 1)模型:

$$\begin{aligned} border_t = & \beta_0 + \beta_1 border_{t-1} + \beta_2 border_{t-2} + \gamma_1 drought_{t-1} \\ & + \gamma_2 diff_t + \gamma_3 age_t + \gamma_4 rival_t + \gamma_5 wall_t + \gamma_6 unified_t + \varepsilon_t \end{aligned} \quad (13.15)$$

其中, 变量 `diff`, `age`, `rival`, `wall` 与 `unified` 被认为只有当期作用, 而气候变量 `drought` 则存在滞后效应。

```
. use border.dta, clear
. tsset decade

. reg border l(1/2).border l.drought diff age
rival wall unified,r
```

Linear regression				Number of obs = 211		
				F(8, 202) = 2040.71		
				Prob > F = 0.0000		
				R-squared = 0.9854		
				Root MSE = 1.0889		
border		Robust				
		Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
border						
	L1.	1.518284	.133108	11.41	0.000	1.255825 1.780744
	L2.	-.5586965	.1278243	-4.37	0.000	-.8107376 -.3066555
drought						
	L1.	-.6333046	.3009855	-2.10	0.037	-1.226781 -.0398281
diff						
		-.0069699	.0028159	-2.48	0.014	-.0125222 -.0014175
age						
		-.0264399	.0123658	-2.14	0.034	-.0508224 -.0020573
rival						
		.34148	.1455019	2.35	0.020	.0545827 .6283772
wall						
		.7339998	.2203202	3.33	0.001	.2995774 1.168422
unified						
		.4078538	.2230656	1.83	0.069	-.0319819 .8476894
_cons						
		.8189595	.472603	1.73	0.085	-.1129085 1.750827

被解释变量 border 的两阶滞后均在 1%水平上显著。

变量 L1.drought 在 5% 水平上显著为负，说明气候越干旱，则游牧民族越会为了生存而进攻中原王朝，从而将游牧边界推向南方。

变量 diff 也在 5% 水平上显著为负，说明中原王朝越早于游牧政权建立(根据王朝周期假说，中原相对更弱)，则北方边界纬度越低。

计算气候冲击对游牧边界的长期效应：假如 drought 永久性增加一单位，即从 0 增加到 1(从年年无灾到年年旱灾)，则中国北方边界纬度将变化 $[\gamma_1/(1-\beta_1-\beta_2)]$ 度。

代入相应系数估计值可得：

$$\begin{aligned} & . \text{dis} = .6333046 / (1 - 1.518284 + .5586965) \\ & - 15.671008 \end{aligned}$$

气候冲击对游牧边界的长期效应为 15.67 度，这是一个很大的效应(从北京到海口的纬度差距约为 20 度)。

13.5 误差修正模型

从经济理论而言，相关的变量之间可能存在长期的均衡关系，而变量的短期变动则是向着长期均衡关系的部分调整。

“误差修正模型”(Error Correction Model, ECM)体现这一思想。

考虑最简单的 ADL(1, 1)模型：

$$y_t = \beta_0 + \beta_1 y_{t-1} + \gamma_1 x_{t-1} + \varepsilon_t \quad (13.16)$$

其中， $|\beta_1| < 1$ ，故为平稳过程。

假设经济理论认为 (y, x) 之间存在长期均衡关系：

$$y = \phi + \theta x \quad (13.17)$$

其中， ϕ 与 θ 为待定参数。

对方程 (13.16) 两边求期望，并令 $y^* = E(y_t) = E(y_{t-1})$ ， $x^* = E(x_t) = E(x_{t-1})$ ，可得

$$y^* = \beta_0 + \beta_1 y^* + \gamma_1 x^* \quad (13.18)$$

整理可得

$$y^* = \frac{\beta_0}{(1-\beta_1)} + \frac{\gamma_1}{(1-\beta_1)} x^* \quad (13.19)$$

由此可知， $\phi = \frac{\beta_0}{1-\beta_1}$ ， $\theta = \frac{\gamma_1}{1-\beta_1}$ 。

其中, $\theta = \frac{\gamma_1}{1-\beta_1}$ 为长期乘数, 衡量当 x 永久性变化一单位时, 将导致 y 的永久性变化幅度。

显然, $\beta_0 = (1-\beta_1)\phi$, $\gamma_1 = (1-\beta_1)\theta$ 。

在方程(13.16)两边同减 y_{t-1} :

$$\Delta y_t = \beta_0 - (1-\beta_1)y_{t-1} + \gamma_1 x_{t-1} + \varepsilon_t \quad (13.20)$$

代入 $\beta_0 = (1-\beta_1)\phi$ 以及 $\gamma_1 = (1-\beta_1)\theta$:

$$\Delta y_t = (1-\beta_1)\phi - (1-\beta_1)y_{t-1} + (1-\beta_1)\theta x_{t-1} + \varepsilon_t \quad (13.21)$$

整理可得

$$\Delta y_t = \underbrace{(\beta_1 - 1)(y_{t-1} - \phi - \theta x_{t-1})}_{\text{error correction}} + \varepsilon_t \quad (13.22)$$

$(y_{t-1} - \phi - \theta x_{t-1})$ 衡量上一期对均衡条件 “ $y = \phi + \theta x$ ” 的偏离(误差), 而 $(\beta_1 - 1)(y_{t-1} - \phi - \theta x_{t-1})$ 为根据上期的误差所作的反向修正, 称为 “误差修正项” (error correction term)。

如果 $(y_{t-1} - \phi - \theta x_{t-1}) > 0$, 即 y_{t-1} 高于其均衡值, 则 $(\beta_1 - 1)(y_{t-1} - \phi - \theta x_{t-1}) < 0$, 故平均而言 $\Delta y_t < 0$, 使下一期更靠近均衡条件。

一般的 ADL 模型都可转换成 ECM 模型。

误差修正模型的经济含义明确, 可分别考察长期效应(长期均衡

关系)与短期效应(误差修正效应)。

13.6 移动平均与 ARMA 模型

另一类时间序列模型为“移动平均过程”(Moving Average Process, 简记 MA)。

记一阶移动平均过程为 MA(1):

$$y_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1} \quad (13.23)$$

其中, $\{\varepsilon_t\}$ 为白噪声, 而 ε_t 的系数被标准化为 1。

由于 y_t 可被看成是白噪声的移动平均, 故名。

考虑 q 阶移动平均过程，记为 $\text{MA}(q)$ ：

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (13.24)$$

假设 $\{\varepsilon_t\}$ 为 iid 且服从正态分布，可进行 MLE 估计。

将 $\text{AR}(p)$ 与 $\text{MA}(q)$ 结合起来，可得 $\text{ARMA}(p, q)$ 模型：

$$y_t = \beta_0 + \beta_1 y_{t-1} + \cdots + \beta_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \quad (13.25)$$

其中， $\{\varepsilon_t\}$ 为白噪声。

对于 $\text{ARMA}(p, q)$ 模型，也可进行 MLE 估计。

对于 $\text{MA}(q)$, 如果 $q \rightarrow \infty$, 可得无穷阶移动平均过程, 记为 $\text{MA}(\infty)$:

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots = \mu + \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j} \quad (13.26)$$

其中, $\theta_0 = 1$ (标准化为 1)。

$\text{MA}(\infty)$ 相当于将 y_t 的决定因素追溯到无穷远的过去。

无穷多个随机变量之和, 能否收敛到某个随机变量?

常用的充分条件是, 序列 $\{\theta_j\}_{j=0}^{\infty}$ 为“绝对值可加总”(Absolutely Summable, 简记 AS), 即 $\sum_{j=0}^{\infty} |\theta_j| < \infty$ (有限)。

在 AS 的条件下, $MA(\infty)$ 有定义。

虽然样本容量 T 通常有限, 无法追溯到无穷远的过去, 但 $MA(\infty)$ 在理论上具有重要意义, 因为 $AR(p)$ 与 $ARMA(p, q)$ 都可写为 $MA(\infty)$ 的形式(参见下文)。

13.7 脉冲响应函数

命题 对于 $y_t = \beta_0 + \beta_1 y_{t-1} + \varepsilon_t$, 假设 $|\beta_1| < 1$, 则此 $AR(1)$ 是 $MA(\infty)$ 。

证明: 反复使用迭代法可得

$$\begin{aligned}
y_t &= \beta_0 + \beta_1 y_{t-1} + \varepsilon_t \\
&= \beta_0 + \beta_1(\beta_0 + \beta_1 y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \\
&= (\beta_0 + \beta_0 \beta_1) + \beta_1^2 y_{t-2} + \beta_1 \varepsilon_{t-1} + \varepsilon_t \\
&= (\beta_0 + \beta_0 \beta_1) + \beta_1^2(\beta_0 + \beta_1 y_{t-3} + \varepsilon_{t-2}) + \beta_1 \varepsilon_{t-1} + \varepsilon_t \\
&= \beta_0(1 + \beta_1 + \beta_1^2) + \beta_1^3 y_{t-3} + \beta_1^2 \varepsilon_{t-2} + \beta_1 \varepsilon_{t-1} + \varepsilon_t \\
&= \dots \\
&= \beta_0(1 + \beta_1 + \beta_1^2 + \dots) + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_1^2 \varepsilon_{t-2} + \beta_1^3 \varepsilon_{t-3} + \dots \\
&= \frac{\beta_0}{1 - \beta_1} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_1^2 \varepsilon_{t-2} + \beta_1^3 \varepsilon_{t-3} + \dots
\end{aligned} \tag{13.27}$$

其中，无穷等比级数之和 $(1 + \beta_1 + \beta_1^2 + \dots)$ 等于 $1/(1 - \beta_1)$ 。

上式为 MA(∞)的形式。

可将平稳的 AR(1)看成是过去所有扰动项的总效应之和，离现在越远的扰动项其影响力呈几何级数递减。

从 AR(1)的MA(∞)表达式可知：

$$IRF(j) \equiv \frac{\partial y_{t+j}}{\partial \varepsilon_t} = \beta_1^j \quad (13.28)$$

$(\partial y_{t+j} / \partial \varepsilon_t)$ 表示，当第 t 期的扰动项 ε_t 变化 1 单位时(而其他期的扰动项均不变)，对相隔 j 期的 y_{t+j} 的影响，称为“动态乘子”(dynamic multiplier)。

动态乘子与绝对时间 t 无关，是时间间隔 j 的函数。

将 $(\partial y_{t+j}/\partial \varepsilon_t)$ 视为 j 的函数，称为“脉冲响应函数” (Impulse Response Function, 简记 IRF)。

它刻画的是 y_{t+j} 对 ε_t 的 1 单位脉冲(impulse)的响应(response)。

将 $(j, \partial y_{t+j}/\partial \varepsilon_t)$ 画图，即可得到对 IRF 的直观认识，称为“脉冲响应图”。

类似地， $AR(p)$ 也是 $MA(\infty)$ 。

更一般地， $ARMA(p, q)$ 也是 $MA(\infty)$ 。

例 以数据集 `gdp_china.dta` 为例, 考察 GDP 对数差分(`dlny`)的自回归模型。

为计算脉冲响应函数(IRF), 将 $AR(p)$ 视为一维的向量自回归 (Vector Autoregression, 简记 VAR, 参见下节), 使用以下命令:

```
varbasic x y z, lags(numlist) irf
```

其中, “`varbasic`” 为估计 VAR 模型的便捷命令, 而 “`x y z`” 为 VAR 模型所包含的变量(此例中只有一个变量)。

选择项 “`lags(numlist)`” 表示滞后阶数, 默认为 “`lags(1 2)`” 或 “`lags(1/2)`”, 即滞后二阶。

选择项 “`irf`” 表示画脉冲响应图。

首先，估计 dlny 的 $\text{AR}(1)$ 模型(为与上文一致，不包括 2013 年的观测值)，并画脉冲响应图(参见图 13.6)。

```
. varbasic dlny if year<2013,lags(1) irf
```

Vector autoregression

Sample: 1980 - 2012	No. of obs	=	33
Log likelihood = 80.96115	AIC	=	-4.785524
FPE = .0004889	HQIC	=	-4.755008
Det(Sigma_ml) = .0004331	SBIC	=	-4.694827

Equation	Parms	RMSE	R-sq	chi2	P>chi2
dlny	2	.021471	0.2879	13.34017	0.0003

	dlny	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
dlny						
	dlny					
	L1.	.5362727	.1468267	3.65	0.000	.2484976 .8240478
	_cons	.0437698	.0143197	3.06	0.002	.0157037 .071836

使用命令 `varbasic` 的估计系数与命令 “`reg dlnngdp 1.dlnngdp`” 完全相同，只是命令 `varbasic` 不提供异方差稳健标准误的选择项(时间序列一般不存在异方差问题)。

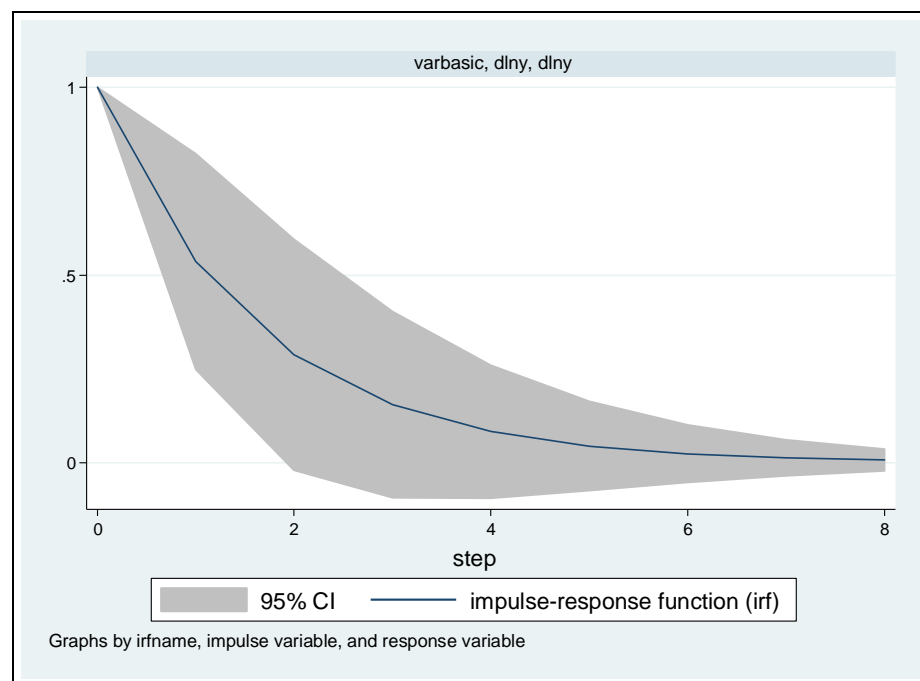


图 13.6 AR(1)模型的脉冲响应函数

AR(1)模型的脉冲响应函数呈指数衰减,从当期的一单位冲击逐渐衰减为 0, 与方程(13.28)的 IRF 表达式一致。

其次, 估计 dlny 的 AR(2)模型, 画 IRF 图(参见图 13.7)。

```
. varbasic dlny if year<2013,irf
```

Vector autoregression

Sample: 1981 - 2012

No. of obs = 32

Log likelihood = 81.69936

AIC = -4.91871

FPE = .0004282

HQIC = -4.873162

Det(Sigma_ml) = .0003548

SBIC = -4.781298

Equation	Parms	RMSE	R-sq	chi2	P>chi2
dlny	3	.019785	0.4234	23.49854	0.0000

dlny	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
dlny						
L1.	.7711595	.1593746	4.84	0.000	.458791	1.083528
L2.	-.4487175	.157661	-2.85	0.004	-.7577275	-.1397076
_cons	.0641134	.0148348	4.32	0.000	.0350377	.0931892

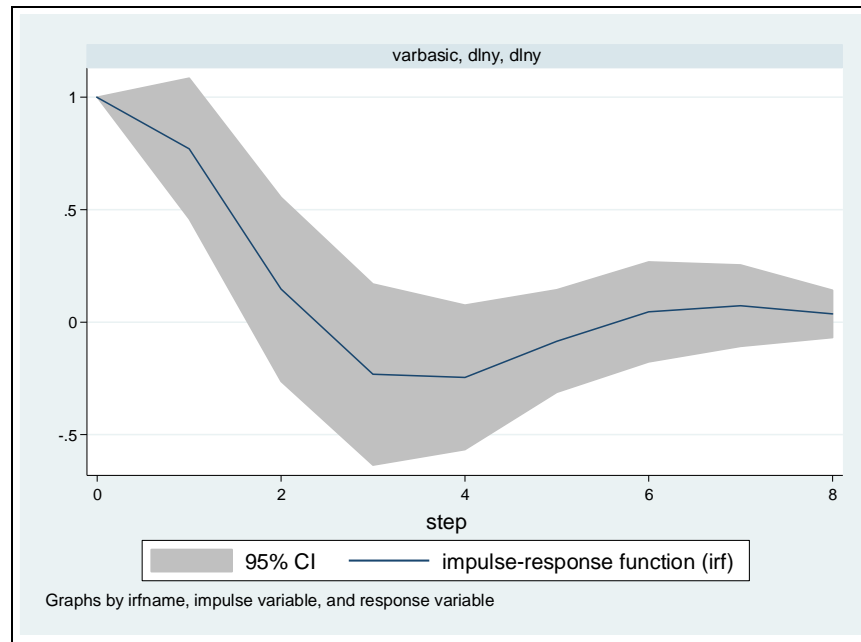


图 13.7 AR(2)模型的脉冲响应函数

AR(2)模型的脉冲响应函数不再单调递减，更具动态特征，先下降，变为负数后再反弹上升，又下降并趋于 0。

13.8 向量自回归过程

常同时关心几个变量的预测，如 GDP 增长率与失业率。

一种方法是用单变量时间序列对每个变量分别作预测。

另一方法将这些变量放在一起，作为一个系统来预测，使得预测相互自洽 (mutually consistent)，称为“多变量时间序列” (multivariate time series)。

Sims(1980)提倡的“向量自回归” (Vector Autoregression, 简记 VAR)正是这样的方法。

假设有两个时间序列 $\{y_{1t}, y_{2t}\}$ ，分别作为两个回归方程的被解释变量。

解释变量为这两个变量的 p 阶滞后值，构成二元的 VAR(p)系统：

$$\begin{cases} y_{1t} = \beta_{10} + \beta_{11}y_{1,t-1} + \cdots + \beta_{1p}y_{1,t-p} + \gamma_{11}y_{2,t-1} + \cdots + \gamma_{1p}y_{2,t-p} + \varepsilon_{1t} \\ y_{2t} = \beta_{20} + \beta_{21}y_{1,t-1} + \cdots + \beta_{2p}y_{1,t-p} + \gamma_{21}y_{2,t-1} + \cdots + \gamma_{2p}y_{2,t-p} + \varepsilon_{2t} \end{cases}$$

(13.29)

其中， $\{\varepsilon_{1t}\}$ 与 $\{\varepsilon_{2t}\}$ 均为白噪声(无自相关)，但允许两个方程的扰动项之间存在“同期相关性”(contemporaneous correlation)：

$$\text{Cov}(\varepsilon_{1t}, \varepsilon_{2s}) = \begin{cases} \sigma_{12} & \text{若 } t = s \\ 0 & \text{其他} \end{cases} \quad (13.30)$$

VAR 的两个方程，其解释变量完全相同。

可以更简洁地写在一起：

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} \\ \beta_{21} \end{pmatrix} y_{1,t-1} + \cdots + \begin{pmatrix} \beta_{1p} \\ \beta_{2p} \end{pmatrix} y_{1,t-p} \\ + \begin{pmatrix} \gamma_{11} \\ \gamma_{21} \end{pmatrix} y_{2,t-1} + \cdots + \begin{pmatrix} \gamma_{1p} \\ \gamma_{2p} \end{pmatrix} y_{2,t-p} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \quad (13.31)$$

将同期变量合成列向量，把相应系数合并为矩阵可得

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \gamma_{11} \\ \beta_{21} & \gamma_{21} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \cdots + \begin{pmatrix} \beta_{1p} & \gamma_{1p} \\ \beta_{2p} & \gamma_{2p} \end{pmatrix} \begin{pmatrix} y_{1,t-p} \\ y_{2,t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \quad (13.32)$$

记 $\mathbf{y}_t \equiv \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}$, $\boldsymbol{\varepsilon}_t \equiv \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$, 则有

$$\mathbf{y}_t = \underbrace{\begin{pmatrix} \beta_{10} \\ \beta_{20} \end{pmatrix}}_{\boldsymbol{\Gamma}_0} + \underbrace{\begin{pmatrix} \beta_{11} & \gamma_{11} \\ \beta_{21} & \gamma_{21} \end{pmatrix}}_{\boldsymbol{\Gamma}_1} \mathbf{y}_{t-1} + \cdots + \underbrace{\begin{pmatrix} \beta_{1p} & \gamma_{1p} \\ \beta_{2p} & \gamma_{2p} \end{pmatrix}}_{\boldsymbol{\Gamma}_p} \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t \quad (13.33)$$

定义相应的系数矩阵为 $\boldsymbol{\Gamma}_0, \boldsymbol{\Gamma}_1, \cdots, \boldsymbol{\Gamma}_p$, 可得

$$\mathbf{y}_t = \boldsymbol{\Gamma}_0 + \boldsymbol{\Gamma}_1 \mathbf{y}_{t-1} + \cdots + \boldsymbol{\Gamma}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t \quad (13.34)$$

此形式与 $\text{AR}(p)$ 相似, 故名 “ $\text{VAR}(p)$ ”。

其中, $\{\boldsymbol{\varepsilon}_t\}$ 是白噪声的推广, 称为“向量白噪声过程”(vector white noise process), 或“新息过程”(innovation process)。

由于 VAR(p) 系统中的解释变量 $\{y_{t-1}, \dots, y_{t-p}\}$ 依赖于 $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots\}$ ，而 ε_t 与 $\{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots\}$ 不相关，故可视所有解释变量为前定变量，与当期扰动项 ε_t 不相关，故可用 OLS 对每个方程分别进行一致估计。

在 VAR 建模时，需确定变量的滞后阶数，及包含几个变量。

滞后阶数的选择

方法一、使用信息准则，比如 AIC 或 BIC。

方法二、检验最后一阶系数的显著性(由大到小的序贯规则)。在上例中，假设要确定使用 VAR(p)还是 VAR($p-1$)，可检验原假设“ $H_0: \beta_{1p} = \beta_{2p} = \gamma_{1p} = \gamma_{2p} = 0$ ”。

方法三、检验 VAR 模型的残差是否为白噪声（是否有自相关）。

如果真实模型为 $\text{VAR}(p)$ ，但被错误设置为 $\text{Var}(p-1)$ ，则解释变量的最后一阶滞后 \mathbf{y}_{t-p} 被纳入扰动项 $\boldsymbol{\varepsilon}_t$ ，导致扰动项出现自相关。

由于 $\{\mathbf{y}_t\}$ 的相关性，包含 \mathbf{y}_{t-p} 的扰动项 $\boldsymbol{\varepsilon}_t$ 将与解释变量 $\{\mathbf{y}_{t-1}, \dots, \mathbf{y}_{t-p+1}\}$ 相关，导致 OLS 估计不一致。

需检验 VAR 模型的残差是否存在自相关。

如果存在自相关，应加入更高阶的滞后变量。

VAR 变量个数的选择

VAR 系统包含的变量个数越多，需要估计的系数越多。

假设有 5 个变量，滞后 4 期，则每个方程中共有 21 个待估系数(含截距项)，整个 VAR 系统共有 105 个待估系数！

待估系数过多使有效样本容量过小，增大估计误差，降低预测精度。故 VAR 模型通常仅包含少数几个变量。

在设定 VAR 模型时，应根据经济理论确定哪些变量在 VAR 模型中。

比如，经济理论告诉我们，通货膨胀率、失业率、短期利息率互相关联，可构成三变量的 VAR 模型。

也可在 VAR 系统中引入其他外生解释变量，比如 $\{w_{1t}, w_{2t}, \dots, w_{Kt}\}$ ，与扰动项不相关。

13.9 VAR 的脉冲响应函数

VAR 模型包含许多参数，其经济意义很难解释，故常将注意力集中于脉冲响应函数。

考虑 n 元 VAR(p) 系统：

$$\mathbf{y}_t = \boldsymbol{\Gamma}_0 + \boldsymbol{\Gamma}_1 \mathbf{y}_{t-1} + \dots + \boldsymbol{\Gamma}_p \mathbf{y}_{t-p} + \boldsymbol{\varepsilon}_t \quad (13.35)$$

其中， \mathbf{y}_t 包含 n 个变量。

正如 $AR(p)$ 可写为 $MA(\infty)$ ，此 $VAR(p)$ 系统也可写成“向量移动平均过程” (Vector Moving Average Process) $VMA(\infty)$ 的形式：

$$\mathbf{y}_t = \boldsymbol{\alpha} + \boldsymbol{\varepsilon}_t + \boldsymbol{\psi}_1 \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\psi}_2 \boldsymbol{\varepsilon}_{t-2} + \cdots = \boldsymbol{\alpha} + \sum_{i=0}^{\infty} \boldsymbol{\psi}_i \boldsymbol{\varepsilon}_{t-i} \quad (13.36)$$

其中， $\boldsymbol{\psi}_0 \equiv \mathbf{I}_n$ ，而 $\boldsymbol{\psi}_j$ 为 n 维方阵。

直观来看， $\boldsymbol{\varepsilon}_{t-1}$ 对 \mathbf{y}_t 的“边际效应”为 $\boldsymbol{\psi}_1$ 。可以证明，

$$\frac{\partial \mathbf{y}_{t+s}}{\partial \boldsymbol{\varepsilon}_t'} = \boldsymbol{\psi}_s \quad (13.37)$$

其中， $(\partial \mathbf{y}_{t+s} / \partial \boldsymbol{\varepsilon}_t')$ 为 n 维列向量 \mathbf{y}_{t+s} 对 n 维行向量 $\boldsymbol{\varepsilon}_t'$ 求偏导数，故得到 $n \times n$ 矩阵 $\boldsymbol{\psi}_s$ 。

假设 $n = 2$ ，则

$$\frac{\partial \mathbf{y}_{t+s}}{\partial \boldsymbol{\varepsilon}'_t} = \boldsymbol{\psi}_s = \begin{pmatrix} \frac{\partial y_{1,t+s}}{\partial \varepsilon_{1t}} & \frac{\partial y_{1,t+s}}{\partial \varepsilon_{2t}} \\ \frac{\partial y_{2,t+s}}{\partial \varepsilon_{1t}} & \frac{\partial y_{2,t+s}}{\partial \varepsilon_{2t}} \end{pmatrix} \quad (13.38)$$

矩阵 $\boldsymbol{\psi}_s$ 是一维情形下相隔 s 期的动态乘子(dynamic multiplier)向多维的推广，其第 i 行、第 j 列元素等于 $(\partial y_{i,t+s} / \partial \varepsilon_{jt})$ 。

它表示，当第 j 个变量在第 t 期的扰动项 ε_{jt} 增加 1 单位时(而其他变量与其他期的扰动项均不变)，对第 i 个变量在第 $(t + s)$ 期的取值 $y_{i,t+s}$ 的影响。

将 $(\partial y_{i,t+s} / \partial \varepsilon_{jt})$ 视为时间间隔 s 的函数, 即“脉冲响应函数”(IRF)。

脉冲响应函数的缺点是, 它假定在计算 $(\partial y_{i,t+s} / \partial \varepsilon_{jt})$ 时, 只让 ε_{jt} 变动, 而所有其他同期扰动项均不变。

此假定只有当扰动项不存在“同期相关”(contemporaneous correlation)时才成立。但现实中, 同期相关普遍存在。

为此, 从扰动项 ε_t 中分离出相互正交的部分, 记为 \mathbf{v}_t 。

新扰动项 \mathbf{v}_t 的各分量正交(不相关), 且方差均被标准化为 1(故变化一单位, 就是变化一个标准差)。

然后计算当 \mathbf{v}_t 中的某分量变动时，对各变量在不同时期的影响，称为“正交化的脉冲响应函数”(Orthogonalized Impulse Response Function, 简记 OIRF)。

但 OIRF 依然有缺点。

首先，正交化冲击(orthogonalized shocks) \mathbf{v}_t 的经济含义不易解释(\mathbf{v}_t 为 $\boldsymbol{\varepsilon}_t$ 中各分量的线性组合)。

其次，OIRF 依赖于变量的次序(order of variables)；如果改变变量次序，可能得到很不相同的结果。

OIRF 虽使得因果关系更清楚，但代价是需对变量起作用的次序作较强的先验假设，而经济理论通常无法对变量次序给出明确的指南。

在实践中，可借助借助格兰杰因果检验确定两个变量之间的排序(参见下文)。

在难以确定变量次序的情况下，可进行稳健性检验，即对于不同的变量排序，分别画正交化脉冲响应图，然后进行比较。

13.10 格兰杰因果检验

经济学中常需确定因果关系究竟是从 x 到 y ，还是从 y 到 x ，抑或双向因果关系。

格兰杰[Granger(1969)]提出了以下检验思想。

首先，原因必然发生于结果之前。

其次，原因包含有关结果的独特信息，对结果具有解释力或预测力。

因此，如果 x 是 y 的因，但 y 不是 x 的因，则 x 的过去值可帮助预测 y 的未来值，而 y 的过去值却不能帮助预测 x 的未来值。

考虑 ADL(p, p)模型：

$$y_t = \gamma + \sum_{m=1}^p \alpha_m y_{t-m} + \sum_{m=1}^p \beta_m x_{t-m} + \varepsilon_t \quad (13.39)$$

滞后阶数 p 可根据“信息准则”或“由大到小的序贯规则”确定。

估计此模型后，检验原假设 “ $H_0 : \beta_1 = \dots = \beta_p = 0$ ”，即 x 的过去值对预测 y 的未来值有无帮助。

如拒绝 H_0 ，称 x 是 y 的“格兰杰因” (Granger cause)。

将回归模型中 x 与 y 的位置互换，可检验 y 是否为 x 的格兰杰因。

实际操作中，常将 (x, y) 构成二元 VAR 系统，使用 Stata 命令 `vargranger` 进行格兰杰因果检验。

格兰杰因果关系并非真正意义上的因果关系。

充其量只是动态相关关系，表明一个变量是否对另一变量有“预测能力” (predictability)。

在某种意义上，它顶多是因果关系的必要条件，而且格兰杰因果关系也可能由第三个变量所引起。

另外，格兰杰因果检验仅适用于平稳序列，或者有协整关系的单位根过程(详见第 14 章)。

对于不存在协整关系的单位根变量，则只能先差分，得到平稳序列后再进行格兰杰因果检验。

13.11 VAR 的 Stata 命令及实例

与 VAR 相关的 Stata 命令包括(假设变量为 x, y, z)

```
varsoc x y z, maxlag(#)
```

此命令用来计算不同滞后期的信息准则，其中“soc”表示 selection-order criteria，“maxlag(#)”表示最大滞后期，默认值为 4。

```
varbasic x y z, lags(numlist) irf
```

这是估计 VAR 模型的便捷命令。选择项“lags(numlist)”表示滞后阶数，默认为“lags(1 2)”或“lags(1/2)”，即滞后二阶。“irf”表示画(未正交化)脉冲响应图，默认为“oirf”(画正交化脉冲响应图)。

估计 VAR 的正式命令为

```
var x y z, lags(numlist) exog(w1 w2)
```

其中，选择项 “lags(numlist)” 表示滞后阶数，默认为 “lags(1/2)”，即滞后二阶。

如果要滞后三阶，可使用选择项 “lags(1/3)”。

选择项 “exog(w1 w2)” 表示在 VAR 模型中引入外生变量 w1,w2。

```
varlmar
```

估计 VAR 后，对残差是否存在自相关进行 *LM* 检验。

`varstable, graph`

估计 VAR 后，通过特征值检验该 VAR 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(参见第 14 章)。选择项 “graph” 表示画出特征值的几何分布图。

`varwle`

估计 VAR 后，对每个方程以及所有方程的各阶系数的联合显著性进行沃尔德检验，其中 “wle” 表示 Wald lag-exclusion statistics。

`vargranger`

估计 VAR 后，进行格兰杰因果检验。

```
irf create irfname, set(filename) step(#)
replace order(varlist)
```

估计 VAR 后，将有关脉冲响应的结果存为 “irfname” (可自行命名)。

选择项 “set(filename)” 表示建立脉冲文件 “filename”，使之成为当前的脉冲文件(make filename active)，并将脉冲结果 “irfname” 存入此脉冲文件 “filename” (若未使用选择项 “set(filename)” 指定脉冲文件，则将脉冲响应结果存入当前的脉冲文件)；

“step(#)” 表示考察截止 # 期的脉冲响应函数，默认为 “step(8)”；“replace” 表示替代已有的同名脉冲响应结果 irfname (如果有)。

一个脉冲文件 “filename” 可存储多个脉冲响应结果 “irfname”。

选择项 “order(varlist)” 指定变量排序，默认使用估计 VAR 时的变量排序计算正交化 IRF。

```
irf graph irf, impulse(varname) response(varname)
```

画脉冲响应图(未正交化)。选择项 “impulse(varname)” 用于指定脉冲变量，“response(varname)” 用来指定反应变量；默认画出所有变量的脉冲响应图。

```
irf graph oirf, impulse(varname)  
response(varname)
```

画正交化的脉冲响应图，选择项含义同上。

如将以上命令中的“irf graph”改为“irf table”，则将相应信息列表而非画图。

```
fcast compute prefix, step(#)
```

估计 VAR 后，计算被解释变量未来#期的预测值，并把预测值赋予被解释变量加上前缀“prefix”(自行确定)的变量名。

```
fcast graph varlist, observed
```

运行命令“fcast compute”后，将变量“varlist”的预测值画图，其中选择项“observed”表示与实际观测值相比较。

以数据集 `macro_swatson.dta` 为例，进行 VAR 估计。

该数据集包含美国 1960 年第 2 季至 2002 年第 1 季的宏观经济季度变量：`inf` 为通货膨胀率，`dinf` 为通货膨胀率的一阶差分，`unem` 为失业率，`quarter` 为季度(时间变量)。

由于通胀率 `inf` 可能不平稳(参见第 14 章)，考虑其一阶差分 `dinf` 与失业率 `unem` 构成的二元 VAR 系统。

首先，打开数据集，看二者的时间趋势(参见图 13.8)。

```
. use macro_swatson.dta, clear  
. tsline dinf unem, lpattern(solid dash)
```

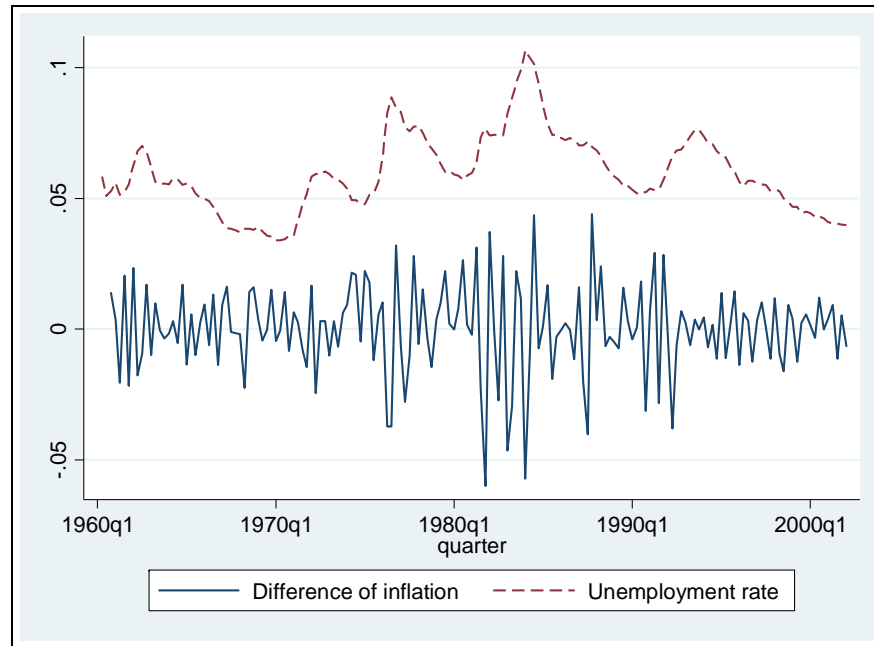


图 13.8 通胀差分与失业率的时间趋势图

其次，根据信息准则，估计此 VAR 系统的阶数。

```
. varsoc dinf unem
```

Selection-order criteria								
Sample: 1961q4 - 2002q1					Number of obs = 162			
lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	881.581				6.6e-08	-10.859	-10.8435	-10.8209
1	1141.4	519.63	4	0.000	2.8e-09	-14.0172	-13.9708	-13.9029
2	1213.86	144.93*	4	0.000	1.2e-09*	-14.8625*	-14.7851*	-14.6719*
3	1217.32	6.9224	4	0.140	1.2e-09	-14.8558	-14.7475	-14.589
4	1219.45	4.2537	4	0.373	1.2e-09	-14.8327	-14.6934	-14.4896

Endogenous: dinf unem
Exogenous: _cons

当 $p = 2$ 时(上表打星号“*”者), AIC 与 BIC 信息准则最小化。

估计二阶向量自回归模型:

```
. var dinf unem, lags(1/2)
```

Vector autoregression

Sample: 1961q2 - 2002q1

No. of obs = 164

Log likelihood = 1224.456

AIC = -14.81044

FPE = 1.27e-09

HQIC = -14.73371

Det(Sigma_ml) = 1.12e-09

SBIC = -14.62143

Equation	Parms	RMSE	R-sq	chi2	P>chi2
dinf	5	.013815	0.3621	93.11047	0.0000
unem	5	.002508	0.9731	5926.554	0.0000

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
dinf						
dinf						
L1.	-.4709436	.0676118	-6.97	0.000	-.6034603	-.338427
L2.	-.401031	.0654944	-6.12	0.000	-.5293976	-.2726644
unem						
L1.	-2.24205	.3371897	-6.65	0.000	-2.90293	-1.58117
L2.	2.03417	.3356694	6.06	0.000	1.37627	2.69207
_cons	.0123764	.0045265	2.73	0.006	.0035047	.0212482
unem						
dinf						
L1.	.0305223	.0122746	2.49	0.013	.0064645	.05458
L2.	-.0113172	.0118902	-0.95	0.341	-.0346216	.0119871
unem						
L1.	1.638031	.0612152	26.76	0.000	1.518051	1.75801
L2.	-.6725807	.0609392	-11.04	0.000	-.7920194	-.5531421
_cons	.0020139	.0008218	2.45	0.014	.0004032	.0036245

大多数系数均很显著。下面检验各阶系数的联合显著性。

. varwle

Equation: dinf			
lag	chi2	df	Prob > chi2
1	72.72164	2	0.000
2	72.9294	2	0.000

Equation: unem			
lag	chi2	df	Prob > chi2
1	741.8916	2	0.000
2	123.1287	2	0.000

Equation: All			
lag	chi2	df	Prob > chi2
1	796.2418	4	0.000
2	187.9991	4	0.000

无论单一方程，还是两个方程作为整体，各阶系数均高度显著。

检验残差是否为白噪声，即残差是否存在自相关。

```
. varlmar
```

Lagrange-multiplier test

lag	chi2	df	Prob > chi2
1	7.3130	4	0.12024
2	5.0645	4	0.28074

H0: no autocorrelation at lag order

可接受残差“无自相关”的原假设。

检验 VAR 系统是否为平稳过程，并画图(参见图 13.9)。

```
. varstable, graph
```


Eigenvalue stability condition	
Eigenvalue	Modulus
.82182	.82182
.7970223	.797022
$-.2258777 + .6292178i$.668533
$-.2258777 - .6292178i$.668533

All the eigenvalues lie inside the unit circle.
VAR satisfies stability condition.

所有特征值均在单位圆之内，故 VAR 模型满足平稳性条件。

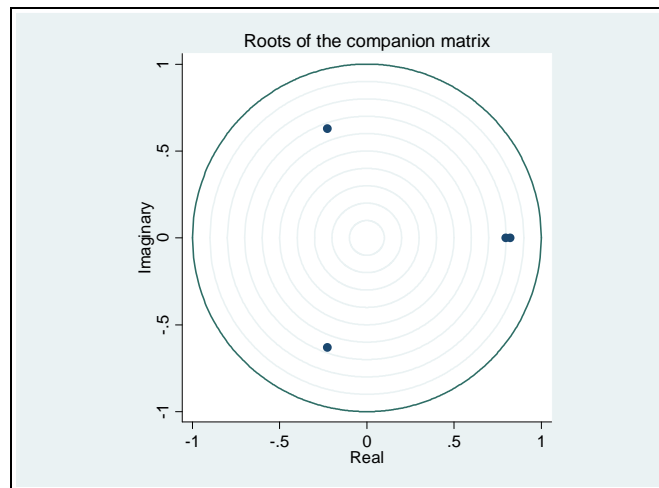


图 13.9 VAR 系统稳定性的判别图

考察变量 `dinf` 与 `unem` 之间的格兰杰因果关系。

. vargranger

Granger causality Wald tests					
Equation	Excluded	chi2	df	Prob >	chi2
dinf	unem	48.123	2	0.000	
dinf	ALL	48.123	2	0.000	
unem	dinf	9.009	2	0.011	
unem	ALL	9.009	2	0.011	

无论以 `dinf` 还是 `unem` 为被解释变量， p 值均远小于 0.05。故二者互为格兰杰原因。格兰杰因果检验无法提供变量排序的信息。

考察正交化脉冲响应函数，将脉冲文件命名为“`macro`”，并将脉冲结果命名为“`iu`” (表示变量排序为 `dinf`, `unem`)。

```
. irf create iu, set(macro)
(file macro.irf created)
(file macro.irf now active)
(file macro.irf updated)
```

此命令建立了脉冲文件“macro.irf”，并将脉冲结果 iu 存入此脉冲文件。

根据此脉冲结果，画正交化的脉冲响应图(结果参见图 13.10)。

```
. irf graph oirf, yline(0)
```

其中，选择项“yline(0)”表示在纵轴 $y=0$ 处画一条水平线。

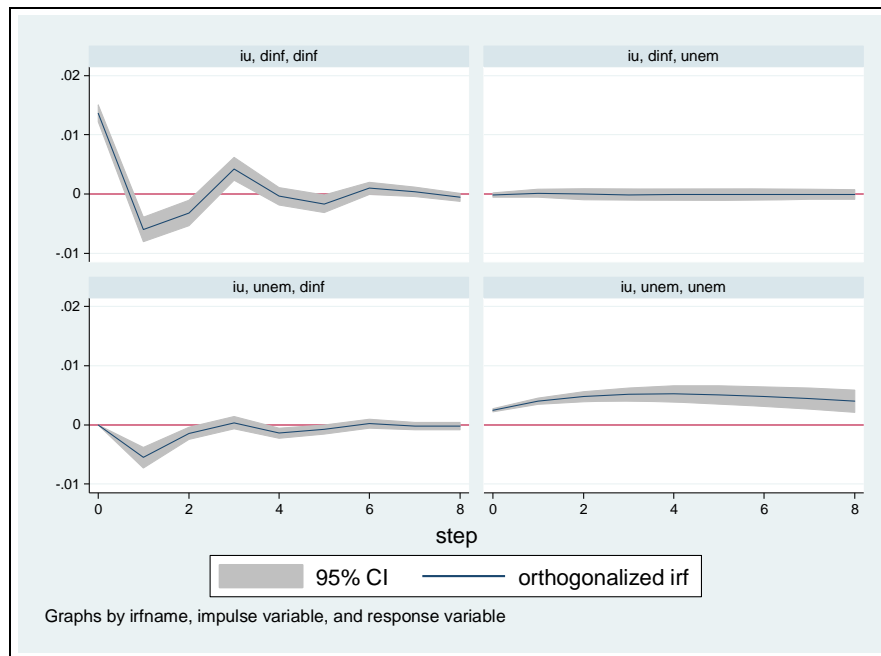


图 13.10 正交化的脉冲响应图

最后一行“Graphs by irfname, impulse variable, and response variable”表明，四个小图的标题命名顺序为“脉冲名称、冲击变量、响应变量”。

比如，左下小图标题为“iu,unem,dinf”，表明此图为根据脉冲结果 iu，冲击变量 unem，响应变量 dinf 所画的脉冲响应图。

它表明，失业率 unem 的一个标准差的正向冲击，将使未来一期的 dinf 下降(根据菲利普斯曲线，失业率上升可缓解通胀压力)，但未来二期的 dinf 即反弹，然后此影响逐渐消失归于 0。

变换变量次序，考察正交化脉冲响应函数的稳健性。

```
. irf create ui, order(unem dinf)
(file macro.irf updated)
```

选择项“order(unem dinf)”表示变量 unem 排在 dinf 之前。此命令将脉冲结果记为 ui。

以上两个脉冲结果 `iu` 与 `ui`，都已存储在当前脉冲文件 `macro.irf` 中；故可直接画图，比较在两种变量排序下，`dinf` 对于 `unem` 冲击的脉冲响应(参见图 13.11)。

```
. irf graph oirf,i(unem) r(dinf) yline(0)
```

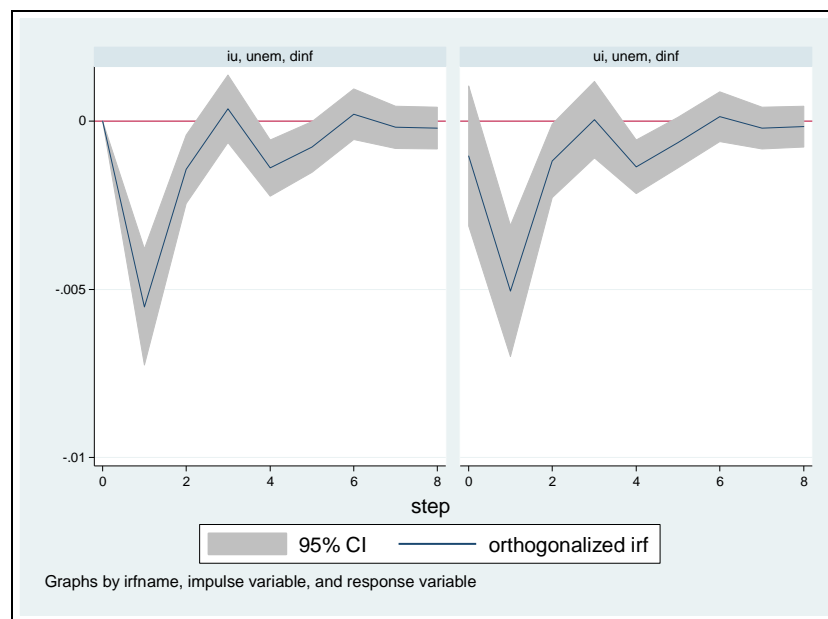


图 13.11 比较两种变量排序下的脉冲响应图(`unem`→`dinf`)

在不同变量排序下，变量 `dinf` 对于 `unem` 冲击的脉冲响应差别不大，只是在反应幅度上略有不同。

比较在两种变量排序下，变量 `unem` 对于 `dinf` 冲击的脉冲响应(参见图 13.12)。

```
. irf graph oirf,i(dinf) r(unem) yline(0)
```

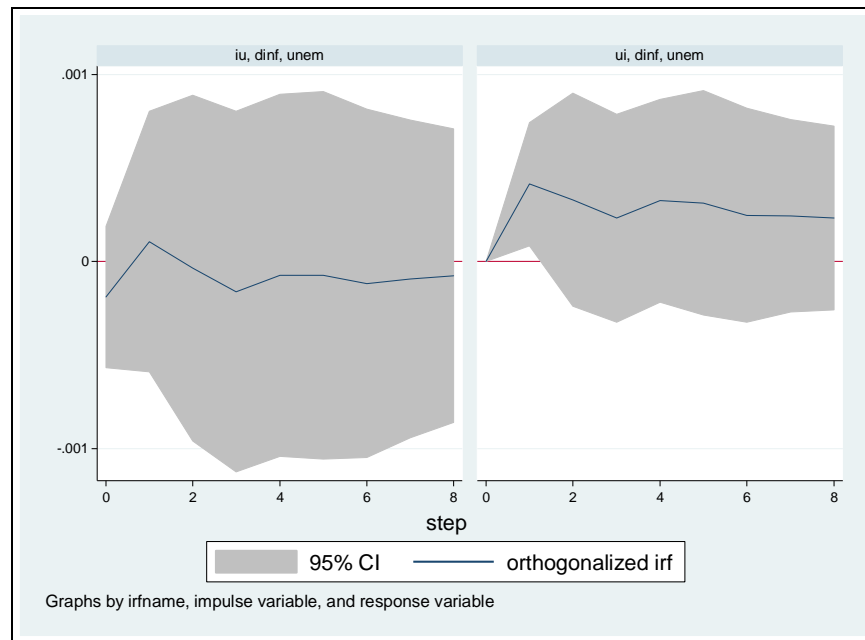


图 13.12 比较两种变量排序下的脉冲响应图(dinf→unem)

变量排序对于从 dinf 至 unem 的脉冲响应幅度有较大影响，但变动方向上依然类似。

估计 VAR 模型后，可用它进行预测。

假设仅用 1999 年以前的数据估计 VAR 模型，然后预测 1999 年 1 季度-2002 年 1 季度的 10 个季度，并与实际观测值比较。

```
. varbasic dinf unem if  
quarter<tq(1999q1),lags(1/2) nograph
```

其中，“tq(1999q1)”表示季度数据格式；选择项“nograph”表示不画脉冲响应图。

Vector autoregression					
Sample: 1961q2 - 1998q4		No. of obs		=	151
Log likelihood = 1118.534		AIC		=	-14.68256
FPE = 1.44e-09		HQIC		=	-14.60139
Det(Sigma_ml) = 1.26e-09		SBIC		=	-14.48274
Equation	Parms	RMSE	R-sq	chi2	P>chi2
dinf	5	.014258	0.3673	87.65165	0.0000
unem	5	.002585	0.9705	4971.522	0.0000
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
dinf					
dinf					
L1.	-.4718995	.0702018	-6.72	0.000	-.6094925 - .3343064
L2.	-.4066509	.0682215	-5.96	0.000	-.5403626 - .2729393
unem					
L1.	-2.279801	.3502631	-6.51	0.000	-2.966304 -1.593298
L2.	2.050438	.3471149	5.91	0.000	1.370105 2.730771
_cons	.0139506	.005028	2.77	0.006	.0040959 .0238053
unem					
dinf					
L1.	.0302607	.012728	2.38	0.017	.0053143 .0552072
L2.	-.0120022	.012369	-0.97	0.332	-.0362449 .0122405
unem					
L1.	1.636296	.0635048	25.77	0.000	1.511829 1.760763
L2.	-.6760615	.062934	-10.74	0.000	-.7994099 -.5527131
_cons	.0023965	.0009116	2.63	0.009	.0006098 .0041832

子样本的样本容量减少为 151。

预测未来 10 个季度的变量取值，分别记为 “f_dinf” 与 “f_unem”。

```
. fcast compute f_,step(10)
```

此命令将生成两个新变量，“f_dinf” 与 “f_unem”，分别为对 dinf 与 unem 的预测值。

这两个预测变量的标准误与置信区间也作为新变量出现在变量窗口。

将 dinf 与 unem 的预测值画图，并与实际观测值比较。

```
. fcast graph f_dinf f_unem, observed  
lpattern(dash)
```

选择项 “observed” 表示显示变量的实际观测值。

选择项 “lpattern(dash)” 表示以虚线来画变量的预测值(以便区别于实际观测值)。

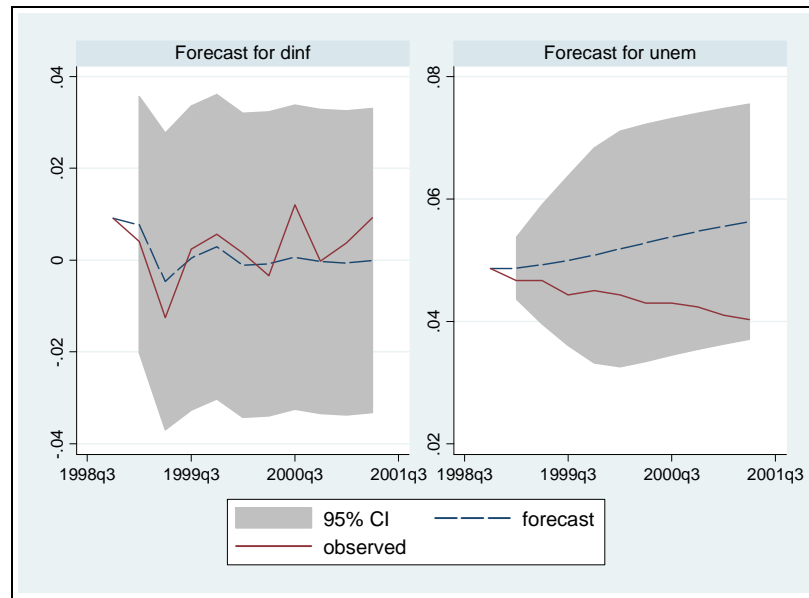


图 13.13 对未来 10 个季度的预测

对通胀率变动的预测准确度优于对失业率的预测。

预测的时期越长，则预测的精确度越低。

13.12 时间趋势项

时间序列常包含某种时间趋势，比如 GDP 的指数增长趋势，或 GDP 对数的线性增长趋势。

如果时间序列 $\{y_t\}$ 包含时间趋势，则不是平稳过程（期望随时间而变）。

一种处理方法为做差分(或者先取对数，再差分)，将其变为平稳过程。

如果想直接对原变量 y_t 建模，可在回归方程中引入“线性时间趋势项” (linear time trend):

$$y_t = \alpha + \beta t + \varepsilon_t \quad (13.40)$$

其中， $t = 1, 2, \dots, T$ 为时间趋势项(T 为样本容量)。

如果 $\{y_t\}$ 存在指数增长趋势，可对其对数建模：

$$\ln y_t = \alpha + \beta t + \varepsilon_t \quad (13.41)$$

其中，系数 β 的经济含义为 y_t 的每期增长率，即

$$\beta = \frac{d \ln y}{dt} = \frac{dy/y}{dt} \quad (13.42)$$

如 y_t 的增长率并非常数，可考虑加入时间趋势的平方项：

$$\ln y_t = \alpha + \beta t + \gamma t^2 + \varepsilon_t \quad (13.43)$$

此时， y_t 的增长率为

$$\frac{dy/y}{dt} = \frac{d \ln y}{dt} = \beta + 2\gamma t \quad (13.44)$$

是否引入时间趋势的平方项，可检验 “ $H_0: \gamma = 0$ ”。

但引入时间趋势的平方项，可能导致多重共线性。

假设样本容量为 100，可在 Stata 中计算时间趋势 t 及其平方项 t^2 的相关系数如下。

```
. set obs 100
. gen t=_n
```



```
. gen t2=t^2  
. corr t t2
```

	t	t2
t	1.0000	
t2	0.9689	1.0000

t 与 t^2 的相关系数接近 0.97，存在严重的多重共线性。

常见的做法是仅包含线性趋势项，以避免多重共线性。

如果 $\{y_t\}$ 包含时间趋势，但被遗漏，可能导致遗漏变量偏差。

考虑以下简单模型：

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad (13.45)$$

其中，变量 y_t 包含时间趋势 γt (但被遗漏)，故可将上式的扰动项写为

$$\varepsilon_t = \gamma t + u_t \quad (13.46)$$

其中， u_t 为不含时间趋势的扰动项。

假设变量 x_t 包含时间趋势 δt ，并可写为

$$x_t = x_t^* + \delta t \quad (13.47)$$

其中， x_t^* 为 x_t 中不包含时间趋势的部分。

将表达式(13.46)与(13.47)代入原模型(13.45):

$$y_t = \alpha + \beta(\underbrace{x_t^* + \delta t}_{x_t}) + (\underbrace{\gamma t + u_t}_{\varepsilon_t}) \quad (13.48)$$

解释变量 x_t 与扰动项 ε_t 相关, 故 OLS 不一致。

由于宏观经济变量通常都有时间趋势, 比如 y_t 与 x_t 都有时间趋势, 故简单地将 y_t 对 x_t 进行回归将发现二者存在显著关系, 而事实上只是因为共同的时间趋势所驱动。

这种现象是“伪回归”(spurious regression)的一种表现。

只要将遗漏的时间趋势加入回归方程(13.45), 即可消除此伪回

归现象：

$$y_t = \alpha + \beta x_t + \gamma t + u_t \quad (13.49)$$

扰动项 u_t 不再包含时间趋势，不会与 x_t 相关，故 OLS 一致。

以数据集 `gdp_china.dta` 为例。直接对 GDP 对数(`lny`)建模，即估计 `lny` 的 AR(2)模型，并加上时间趋势项。

```
. use gdp_china.dta,clear  
  
. gen t=_n  
  
. reg lny l(1/2).lny t if year<2013,r
```

Linear regression					Number of obs = 33	
					F(3, 29) =38047.85	
					Prob > F = 0.0000	
					R-squared = 0.9997	
					Root MSE = .01627	
lny	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
lny						
L1.	1.302829	.1150665	11.32	0.000	1.067491	1.538166
L2.	-.7357654	.0967217	-7.61	0.000	-.9335835	-.5379473
t	.0411649	.0092175	4.47	0.000	.022313	.0600168
_cons	3.473641	.7552359	4.60	0.000	1.929011	5.018272

lny 的两阶滞后以及时间趋势项都高度显著。

预测 GDP 对数，并记为 lny3，然后计算 2013 年的预测误差。

```
. predict lny3
(option xb assumed; fitted values)
(2 missing values generated)
```

```
. dis exp(lny3[36])  
95597.887
```

```
. dis y[36]-exp(lny3[36])  
-508.67625
```

“AR(2)+时间趋势项”模型的预测误差为-508.67625 亿元，即高估了 508.67625 亿元。

dlny 的 AR(2)模型的预测高估了 680.78688 亿元(参见上文)。

在此例，直接对 GDP 对数(lny)建模并引入时间趋势项，其预测能力优于对 GDP 对数差分(dlny)的平稳过程建模。

13.13 季节调整

1. 季节效应

对于月度或季度时间序列，常需进行“季节调整”(seasonal adjustment)，去掉“季节效应”后才能使用。

例 考察中国的季度 GDP 数据。由于第一季度包含春节，故通常第一季度的 GDP 偏低。

如果直接以第二季度 GDP 除以第一季度 GDP 来计算环比增长率，则会高估第二季度的 GDP 增长率；将第一季度 GDP 除以上年第四季度 GDP 则会低估第一季度的 GDP 增长率。

包含季节效应的时间序列不能直接计算环比增长率。

如不进行季节调整，只能计算同比增长率，即与去年同一季(月)相比。

对于年度数据，则不需要进行季节调整。

可能导致季节效应的因素包括：

- (1) 天气因素：比如，在冬季由于取暖而增加能源消耗。
- (2) 行政因素：比如，学校开学与放假对交通的影响。

(3) 固定假日：比如，十一国庆节对旅游与交通的影响。

(4) 移动假日(moving holiday)：比如，春节期间，铁路运输量增加而 GDP 下降。

(5) 日历因素：比如，闰年与闰月的影响。

(6) 交易日效应：比如，五金店销售额在有五个周末的月份高于只有四个周末的月份。

所有这些季节因素共同构成一个时间序列的“季节要素”(seasonal component)。

该时间序列的长期走势与中期周期称为“趋势循环要素”(trend cycle component), 简称“趋势要素”(trend component)。

其他不可预测的随机扰动为该序列的“不规则要素”(irregular component)。

2. 季节调整的原理

季节调整通过估计“季节因子”(seasonal factor)来进行。

根据季节因子起作用的方式, 季节因子分为两种, 即“加法季节因子”(additive seasonal factor)与“乘法季节因子”(multiplicative seasonal factor)。

“加法季节因子”意味着对所有第 1 月(或第 1 季)都加上相同的季节因子，以此类推。

“加法模型”(additive model)的数学表达式如下：

$$Y_t = TC_t + S_t + I_t \quad (13.50)$$

其中， Y_t 为原序列， TC_t 为趋势循环要素， S_t 为季节要素，而 I_t 为不规则要素。

“乘法季节因子”意味着对所有第 1 月(或第 1 季)都乘以相同的季节因子，以此类推。

“乘法模型” (multiplicative model)的数学表达式如下：

$$Y_t = TC_t \times S_t \times I_t \quad (13.51)$$

使用乘法模型要求 Y_t 序列中不包含零或负数。对方程两边取对数：

$$\ln Y_t = \ln TC_t + \ln S_t + \ln I_t \quad (13.52)$$

方程(13.52)在形式上与加法模型相同，称为“对数加法模型”。

季节调整的目标是将原序列 Y_t 分解为趋势循环要素、季节要素与不规则要素，然后去掉季节要素 S_t ，得到季节调整序列 (seasonally adjusted series)，即趋势循环要素与不规则要素之和。

季节调整的方法有多种，使用不同方法，会得到不同的季节调整序列，带有一定的主观性；这是季节调整的局限性。

下面介绍最为简便的回归法。

3. 回归法

首先生成月度(或季度)虚拟变量，然后把时间序列对这些虚拟变量进行 OLS 回归，所得残差就是经季节调整后的序列。

以 `airpassengers.dta` 为例，该数据集包括 1949-1960 年国际航空旅客人数的月度数据(`airpassengers`)与时间变量(`time`)。

打开数据集后，看 airpassengers 的时间趋势图。

```
. use airpassengers.dta,clear
```

```
. tsset time
```

```
time variable:  month, 1949m1 to 1960m12  
               delta:  1 month
```

```
. tsline airpassengers
```

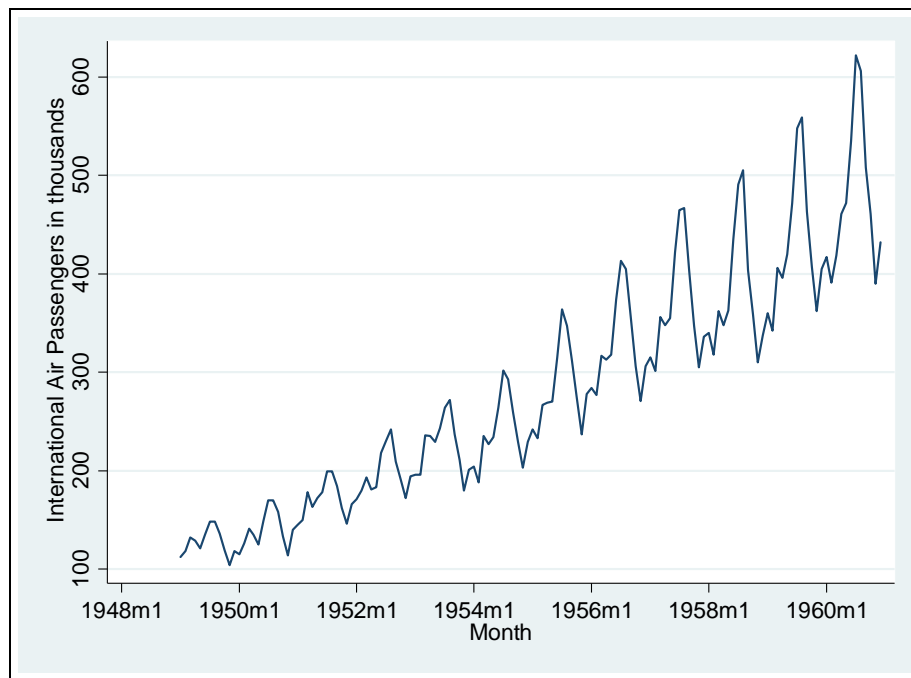


图 13.14 国际航空旅客人数的季节波动

国际航空旅客人数存在明显的季节波动，在夏季(七、八两月)达到高峰。

为生成月度虚拟变量，首先从时间变量 `time` 提取月度信息，记为变量 `month`。

```
. gen month=month((dofm(time)))
```

此命令将生成变量 `month`，取值为 `1, 2, ..., 12`，对应于一年中的 12 个月。

其次，使用命令 `tab` 生成月度虚拟变量。

```
. tab month, gen(m)
```

其中，选择项 “`gen(m)`” 表示，根据变量 `month` 的不同取值，生成相应的虚拟变量，记为 `m1, m2, ..., m12`，分别对应于 12 个月。

month	Freq.	Percent	Cum.
1	12	8.33	8.33
2	12	8.33	16.67
3	12	8.33	25.00
4	12	8.33	33.33
5	12	8.33	41.67
6	12	8.33	50.00
7	12	8.33	58.33
8	12	8.33	66.67
9	12	8.33	75.00
10	12	8.33	83.33
11	12	8.33	91.67
12	12	8.33	100.00
Total	144	100.00	

以 1 月份为参照值，把变量 `airpassengers` 对第 2-12 月的月度虚拟变量进行回归。

```
. reg airpassengers m2-m12
```

Source	SS	df	MS	Number of obs = 144		
Model	218382.243	11	19852.9312	F(11, 132) = 1.42		
Residual	1839661.92	132	13936.8327	Prob > F = 0.1690		
Total	2058044.16	143	14391.9172	R-squared = 0.1061		
				Adj R-squared = 0.0316		
				Root MSE = 118.05		
airpasseng~s	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
m2	-6.75	48.19549	-0.14	0.889	-102.0854	88.58545
m3	28.41667	48.19549	0.59	0.556	-66.91878	123.7521
m4	25.33333	48.19549	0.53	0.600	-70.00211	120.6688
m5	30.08333	48.19549	0.62	0.534	-65.25211	125.4188
m6	69.91667	48.19549	1.45	0.149	-25.41878	165.2521
m7	109.5833	48.19549	2.27	0.025	14.24789	204.9188
m8	109.3333	48.19549	2.27	0.025	13.99789	204.6688
m9	60.66667	48.19549	1.26	0.210	-34.66878	156.0021
m10	24.83333	48.19549	0.52	0.607	-70.50211	120.1688
m11	-8.916667	48.19549	-0.19	0.854	-104.2521	86.41878
m12	20.08333	48.19549	0.42	0.678	-75.25211	115.4188
_cons	241.75	34.07936	7.09	0.000	174.3377	309.1623

七、八两月的虚拟变量(m7 与 m8)均在 5%水平上显著为正；其他月份的虚拟变量不显著。

为得到经季度调整的序列，使用命令 `predict` 计算上述回归的残差项(记为 `air_sa`)。

```
. predict air_sa,r
```

但 OLS 残差项的平均值一定为 0, 故需把原序列的均值加回去, 并记季节调整序列为 `airpassengers_sa`。

```
. sum airpassengers
```

Variable	Obs	Mean	Std. Dev.	Min	Max
airpasseng~s	144	280.2986	119.9663	104	622

```
. gen airpassengers_sa = air_sa+r(mean)
```

将回归法季节调整序列与原序列画图，参见图 13.15。

```
. tsline airpassengers_sa  
airpassengers, lpattern(dash)
```

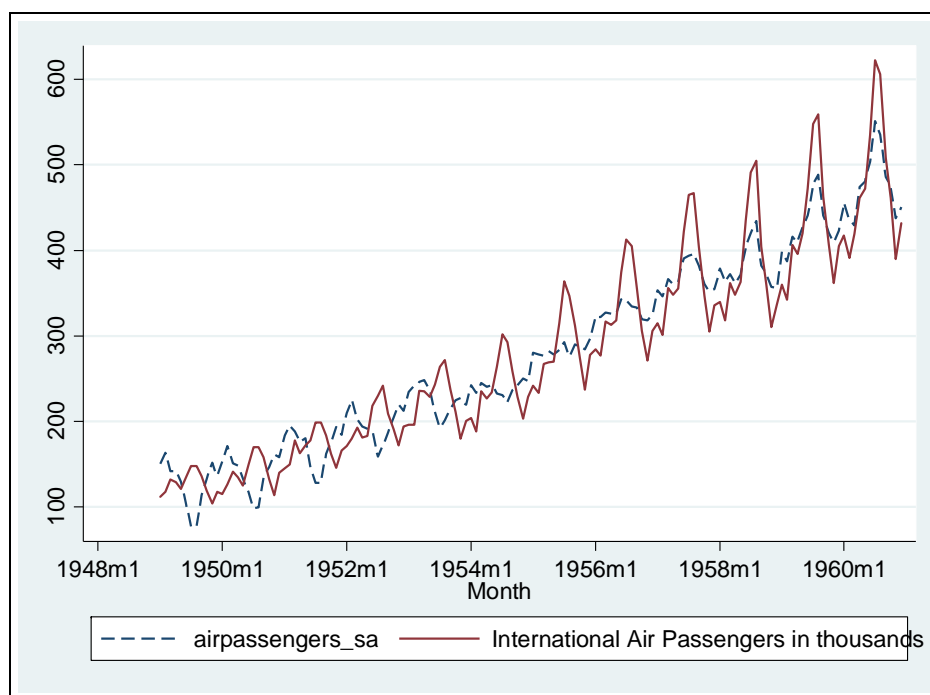


图 13.15 季节调整序列与原序列

13.14 日期数据的导入

如果时间序列为年度数据,将时间数据导入 Stata 并无特殊之处,只要使用命令 “`tsset year`” 将变量 `year` 设为时间变量即可(假设时间变量为 `year`)。

对于日度、月度及季度数据,导入方法略为复杂。

如果数据中含有格式为 “1949-10-01” 或 “1949/10/01” 的时间变量,在导入 Stata 后,可能被视为 “字符串” (string),而非 “数字型” (numeric),无法直接对其进行运算(比如,滞后一期)。

对于日度数据(daily data),可使用如下命令转换为 “整数日期变量” (integer date variable):

```
gen newvar=date(varname,"YMD")
```

其中，函数“date”表示转换为日期变量；“varname”为原来的时间变量，“newvar”为新定义的时间变量。

"YMD"告诉 Stata，原始数据的格式为“年-月-日”。如数据格式为“月-日-年”，则应为"MDY"，以此类推。

如此定义后，新的时间变量 newvar 将以“整数日期”的形式显示。

在 Stata 内部，所有日期变量的存储格式均为“elapsed dates” (Stata 称为 Stata Internal Format，简记 SIF)，即从 1960 年 1 月 1 日以来过了多少天。

为让新的时间变量 `newvar` 仍以通常的日期格式(Human Readable Format, 简记 HRF)在 Stata 中显示, 可输入命令:

```
format newvar %td
```

其中, “%td” 中的 “d” 即表示 “date”。

对于月度数据(monthly data), 可使用如下命令进行转换

```
gen newvar=monthly(varname, "YM" )
```

其中, 函数 “monthly” 表示转换为月度变量; “YM” 告诉 Stata, 原始数据的格式为 “年-月”。

对于新定义的时间变量 `newvar`，Stata 内部的月度变量存储格式为 “`elapsed months`”，即从 1960 年 1 月以来过了多少月。

如仍想以日期格式在 Stata 中显示，可输入命令：

```
format newvar %tm
```

其中，“`%tm`” 中的 “`m`” 即表示 “month”。

对于季度数据(quarterly data)，可使用以下命令进行转换

```
gen newvar=quarterly(varname, "YQ")
```

其中，函数 “`quarterly`” 表示转换为季度变量；“`YQ`” 告诉 Stata，原始数据的格式为 “年-季”。

对于新定义的时间变量 `newvar`，Stata 内部的季度变量存储格式为 “elapsed quarters”，即从 1960 年第 1 季以来过了多少季度。

如仍想以日期格式在 Stata 中显示，可输入命令：

```
format newvar %tq
```

其中，“%tq” 中的 “q” 即表示 “quarter”。

如在原始数据中，年、月、日分别以数字型(numeric)变量 “Y, M, D” 来表示，可使用以下命令将其合成单一的日期变量：

```
gen newvar=mdy(M,D,Y)
```

有关日期数据的更多更多说明，参见 “help date”。