

第七讲：非平稳序列：差分平稳及 ARIMA 模型

概要： 本讲介绍均值非平稳序列分析。非平稳时间序列分析的理论基础是 Cramer 分解定理，即差分可有效提取确定性非平稳信息。在介绍差分平稳概念的基础上，给出了理解 ARIMA 模型的思路、模型的统计性质、建模步骤。

一、Cramer 分解定理

Wold 定理是现代时间序列分析的灵魂，1961 年，Wold 的老师 Cramer 证明 Wold 分解定理的分解思路同样可以用于非平稳序列。即，任意一个时间序列 $\{X_t\}$ 都可以分解成两个不相关的时间序列之和，一个是纯确定性 purely deterministic，另一个是纯随机性 purely stochastic，可以写成：

$$X_t = V_t + \xi_t \quad 7.1$$

式中， V_t 是纯确定性序列， ξ_t 为纯随机序列。Cramer 定理还证明，一个纯随机过程 purely stochastic process 能够写成一系列白噪声过程 $\{\varepsilon_t\}$ 的组合，一个纯确定性过程 purely deterministic process 能够写成由时间 t 的多项式，即：

$$\xi_t = \Psi(B) \varepsilon_t \quad 7.2$$

$$V_t = \sum_{j=0}^d \beta_j t^j \quad 7.3$$

$$X_t = V_t + \xi_t = \sum_{j=0}^d \beta_j t^j + \Psi(B) \varepsilon_t \quad 7.4$$

式中， $d < \infty$ ； $\beta_1, \beta_2, \dots, \beta_d$ 为常数系数； $\{\varepsilon_t\}$ 为一个零均值白噪声序列； B 为延迟算子。

因为

$$E(\xi_t) = \Psi(B)E(\varepsilon_t) = 0 \quad 7.4$$

所以

$$E(X_t) = E(V_t + \xi_t) = E(V) = \sum_{j=0}^d \beta_j t^j \quad 7.5$$

即，式 7.5 表明 $\{X_t\}$ 受到确定性的影响，其均值为时间 t 的函数，式 7.4 反映了 $\{X_t\}$ 受到随机性的影响。

Cramer 定理说明任何一个序列的波动都可以视为同时受到确定性影响和随机性影响的作用。平稳序列要求这两个方面都是稳定的，都不随时间 t 的变化而变化，而非平稳序列产生的机理就在于它所受到的这两方面的影响至少有一个方面是不稳定的，在本讲中，不稳定是由确定性部分产生的。

二、差分平稳 Difference Stationary

2.1 差分运算的实质

拿到观测值序列之后，分析的重点就是通过有效的手段提取序列中所蕴涵的确定性信息。确定性信息提取的方法非常多，其中使用最为广泛的是差分方法。Crammer 分解定理在理论上保证了适当阶数的差分一定可以充分提取确定性信息。

Cramer 分解定理告诉我们，非平稳序列的确定性信息构成如下：

$$E(X_t) = E(V_t + \xi_t) = E(V) = \sum_{j=0}^d \beta_j t^j \quad 7.5$$

显然，在 Cramer 分解定理的保证下， d 阶差分就可以将 $\{X_t\}$ 中蕴含的确定性信息提取出来，即：

$$\nabla^d V_t = \nabla^d \sum_{j=0}^d \beta_j t^j = c, \quad c \text{ 为一常数} \quad 7.6$$

微分对象是连续型变量，而差分对象是离散型变量。如果式 7.6 是连续型变量，求导：

$$\frac{dV_t}{d^d t} = \frac{d(\beta_0 + \beta_1 t + \beta_2 t^2 + \cdots + \beta_d t^d)}{d^d t} = \beta_d$$

从中可以理解式 7.6 的结果。

依据差分的定义，有

$$\nabla X_t = X_t - X_{t-1}$$

等价于

$$X_t = X_{t-1} + \nabla X_t \quad 7.7$$

这意味着 1 阶差分的实质就是一个 1 阶自回归过程，它是用延迟一期的历史过程 $\{x_{t-1}\}$ 作为自变量来解释当期序列 $\{x_t\}$ 的变动情况，差分序列 $\{\nabla x_t\}$ 度量的是 1 阶自回归过程中产生的随机误差的大小。

展开 2 阶、3 阶差分，分别为：

$$\nabla^2 X_t = \nabla X_t - \nabla X_{t-1} = X_t - 2X_{t-1} + X_{t-2} = (1 - 2B + B^2)X_t = (1 - B)^2 X_t$$

$$\nabla^3 X_t = \nabla^2 X_t - \nabla^2 X_{t-1} = X_t - 3X_{t-1} + 3X_{t-2} - X_{t-3} = (1 - 3B + 3B^2 - B^3)X_t = (1 - B)^3 X_t$$

依次类推，得出：

$$\nabla^d X_t = (1 - B)^d X_t = \sum_{i=0}^d (-1)^i C_d^i X_{t-i} = X_t - \sum_{i=1}^d (-1)^{i+1} C_d^i X_{t-i} \quad 7.8$$

式 7.7 移项，展开得：

$$X_t = \sum_{i=1}^d (-1)^{i+1} C_d^i X_{t-i} + \nabla^d X_t \quad 7.9$$

比较式 7.6、式 7.8 和式 7.9 可知，差分运算得实质就是使用自回归的方式提取序列中蕴含得确定性信息，这里 $\{\nabla^d X_t\}$ 度量得是 d 阶自回归过程中产生的随机误差的大小。

2.2 差分方式的选择

实践中，可以依据序列的不同特点选择合适的差分方式，常见的有以下四种：

1、序列蕴含显著的线性趋势，1 阶差分就可以实施趋势平稳。

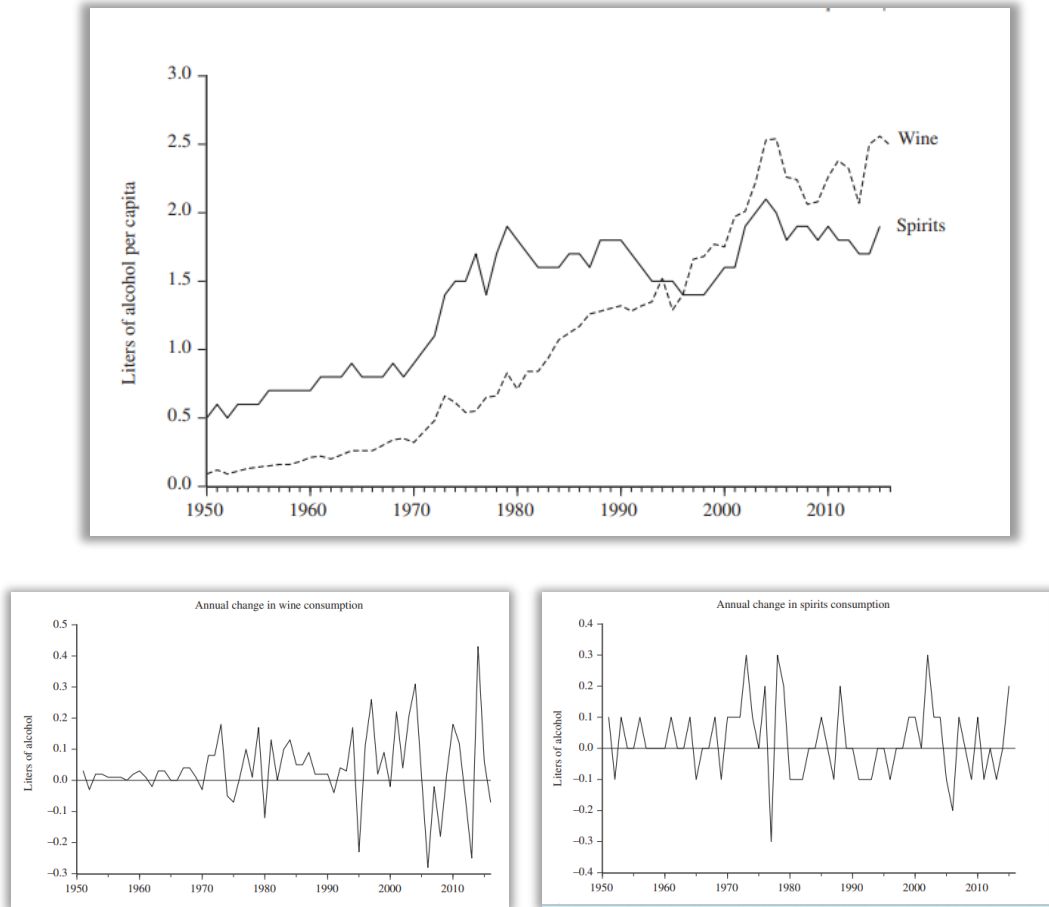


图 7.1 时序图（上）和 1 阶差分图（下）

图 7.1 的上图是时序图，给出了 1950 年-2015 年人均消费酒精性饮料的数量（包括葡萄酒 Wine 和烈性酒 Spirits），均呈现向右上方倾斜的趋势。图 7.1 的下图是 1 阶差分图，左边是人均葡萄酒消费的逐年变动图，右边是人均烈性酒消费的逐年变动图。可以看出，趋势性信息已被 1 阶差分消除，即 $\nabla X_t = X_t - X_{t-1}$ 。

2、序列蕴含曲线趋势，通常 2 阶差分就可以提取曲线趋势信息

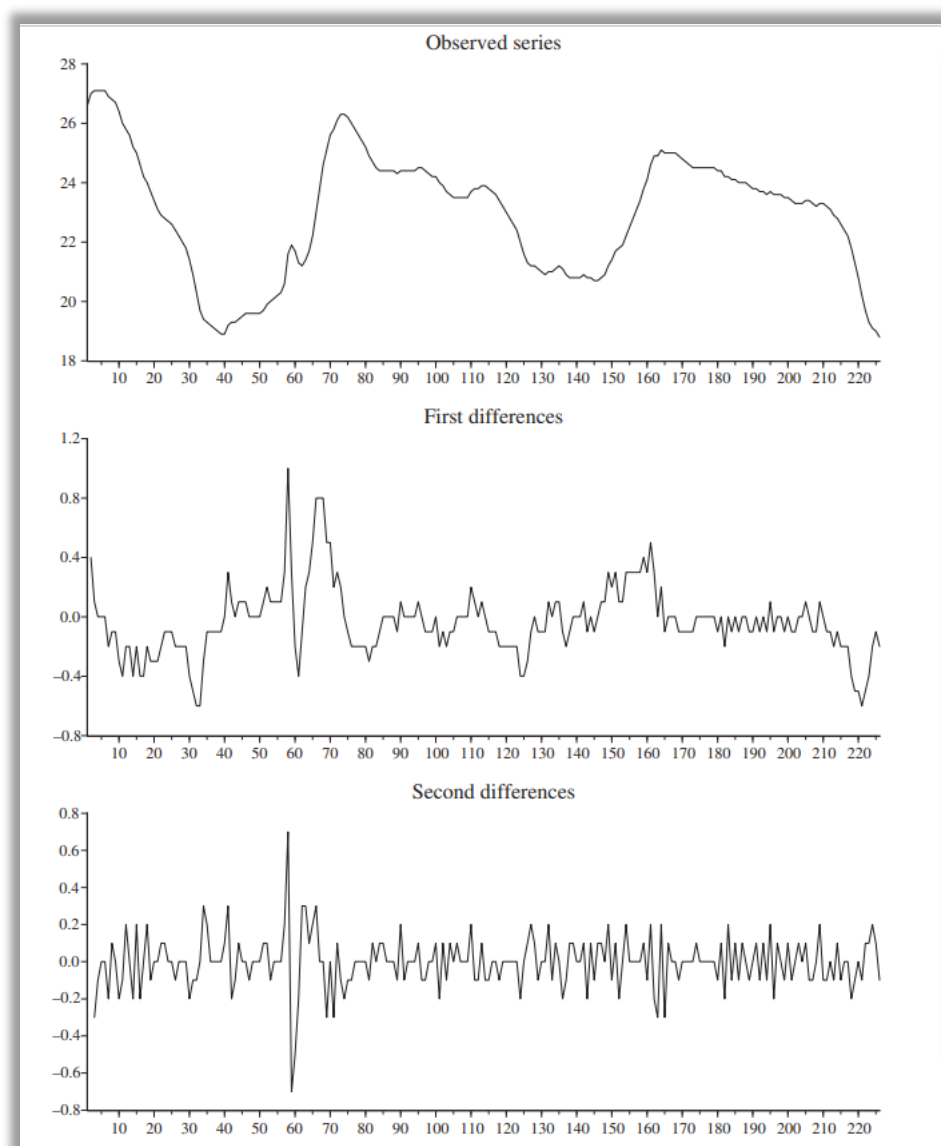


图 7.2 时序图（上）、1 阶差分图（中）和 2 阶差分图（下）

在某些情况下，1 阶差分不足以形成平稳时序，则需要更高阶的差分（通常是 2 阶或 3 阶）。图 7.2 的时序图（上）显示的是某化学反应的连续温度读数，呈现出比较独特的非平稳形式，存在着随机的趋势以及水平量的变动。虽然 1 阶差分（中）弱化了趋势和变动，但并未完全清除，因此，需要进行 2 阶差分（下），即

$$\nabla^2 X_t = \nabla X_t - \nabla X_{t-1} = X_t - 2X_{t-1} + X_{t-2}$$

3、实际应用的要求

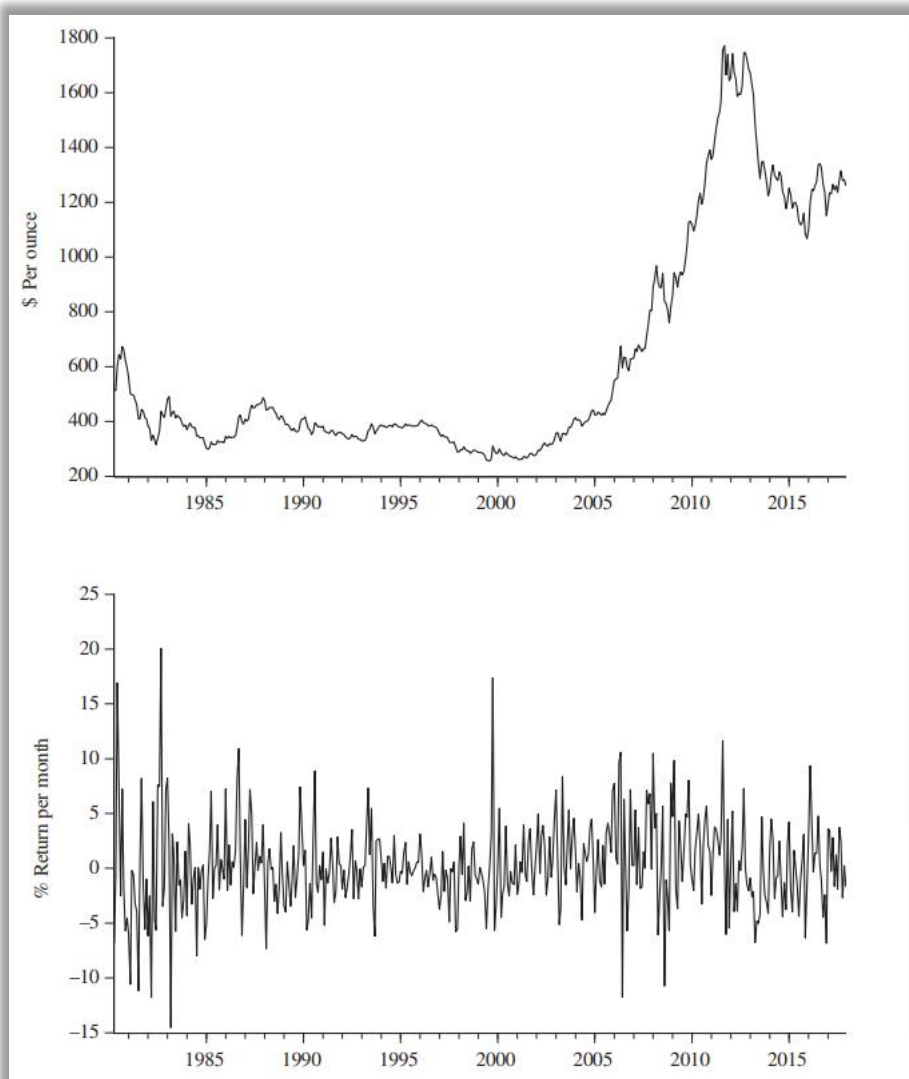


图 7.3 1980 年-2017 年黄金月度价格变动图（上）及月度回报率（下）

对于某些时间序列，为了便于经济解释，通常不是采用简单的差分，而是采用发展速度 $\nabla X_t/X_{t-1}$ 或 $100\nabla X_t/X_{t-1}$ 进行变换。对于金融时间序列，这类变换被定义成回报率 **Return**。图 7.3 描绘了 1980 年-2017 年黄金的月度价格及其回报率（ $100\nabla X_t/X_{t-1}$ ）。价格显然是非平稳的，主要的变化从 2001 年的夏天 270 美元/盎司上升到 2012 年秋天的 1750 美元/盎司，但月度回报率（ $100\nabla X_t/X_{t-1}$ ）虽然波动剧烈（存在异方差），但不存在趋势。

为了简化 $\nabla X_t/X_{t-1}$ 及 $100\nabla X_t/X_{t-1}$ 变换的过程，有以下简要公式：

$$\frac{\nabla X_t}{X_{t-1}} = \frac{X_t - X_{t-1}}{X_{t-1}} = \frac{X_t}{X_{t-1}} - 1 \approx \ln\left(\frac{X_t}{X_{t-1}}\right) = \ln X_t - \ln X_{t-1} = \nabla \ln X_t \quad 7.10$$

即，先对观测值序列 $\{x_t\}$ 取对数，形成 $\{\ln x_t\}$ 序列，再进行 1 阶差分。

还有其他的依据实际序列的变换方法，限于时间，这里不深入介绍。

4、蕴涵固定周期的序列

对于蕴含固定周期的序列进行步长为周期长度的差分运算，通常可以较好地提取周期信息。图 7.4 给出了 1962 年 1 月至 1975 年 12 月平均每头奶牛月产奶量序列的时序图，从中可以看到，该序列具有一个线性递增的长期趋势和一个周期长度为一年的季节变动（季节长度为 1 月，一年 12 季）。

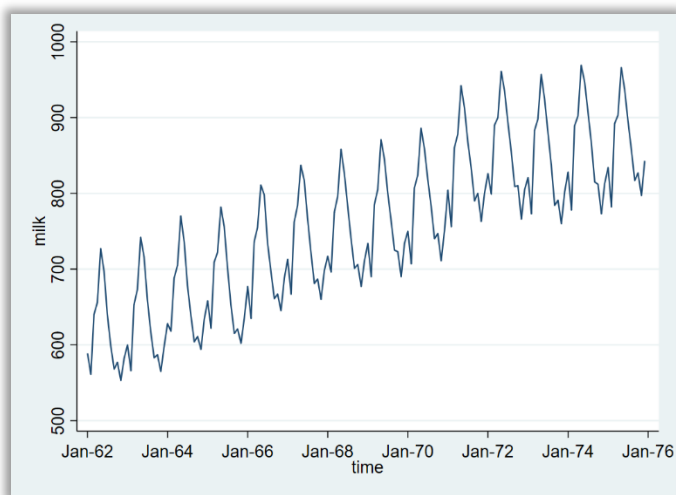


图 7.4 奶牛月产量序列时序图

对原序列先做一阶差分，提取线性递增趋势：

$$\nabla X_t = X_t - X_{t-1}$$

一阶差分后序列 $\{\nabla X_t\}$ 的时序图如图 7.5 所示。

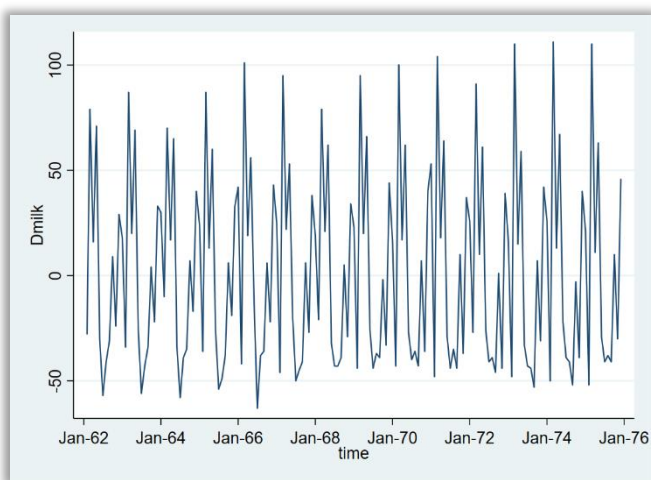


图 7.5 奶牛月产量 1 阶差分后序列时序图

图 7.5 显示，1 阶差分后线性递增信息被提取，但序列还具有稳定的季节波动和随机波动，由此，对 1 阶差分后序列再进行 12 步差分，提取季节波动信息。

$$\nabla_{12}\nabla X_t = \nabla X_t - \nabla X_{t-12} = X_t - X_{t-1} - (X_{t-12} - X_{t-13})$$

周期差分后的序列 $\{\nabla_{12}\nabla X_t\}$ 的时序图如图 7.6 所示。

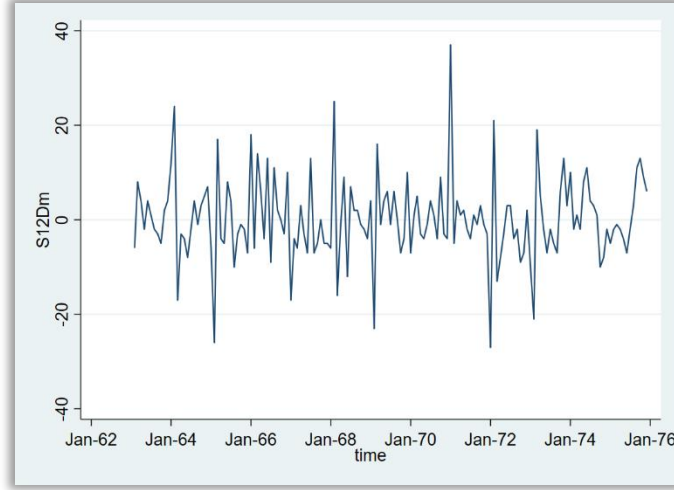


图 7.6 奶牛月产量 1 阶差分 12 步差分后序列时序图

图 7.6 显示，周期差分可以非常好地提取周期信息，而差分后的平稳序列呈现典型的随机波动特征，说明原序列中蕴涵的长期趋势和季节效应等确定性信息已经较好地提取。

2.3 过差分

足够多次的差分运算可以充分提取原序列中的非平稳确定性信息，但差分运算的阶数并不是越多越好。因为差分运算时一种对信息的提取、加工过程，每次差分都会有信息的损失，所以在实际应用中差分运算的阶数要适当，应当避免出现过度差分现象。

假定线性非平稳序列 $\{x_t\}$ 形如：

$$X_t = \beta_0 + \beta_1 t + \varepsilon_t \quad 7.11$$

式中， $E(\varepsilon_t) = 0$ ， $Var(\varepsilon_t) = \sigma_\varepsilon^2$ ， $Cov(\varepsilon_t, \varepsilon_{t-1}) = 0$ ， $t \geq 1$ ，

对 x_t 做 1 阶差分：

$$\nabla X_t = X_t - X_{t-1} = \beta_0 + \beta_1 t + \varepsilon_t - [\beta_0 + \beta_1(t-1) + \varepsilon_{t-1}] = \beta_1 + \varepsilon_t - \varepsilon_{t-1} \quad 7.12$$

显然，1 阶差分后序列 $\{\nabla X_t\}$ 为平稳序列。这说明 1 阶差分运算有效地提取了 $\{x_t\}$ 中的非平稳的确定性信息。对 1 阶差分后的序列 $\{\nabla X_t\}$ 再做一次差分：

$$\nabla^2 X_t = \nabla X_t - \nabla X_{t-1} = \beta_1 + \varepsilon_t - \varepsilon_{t-1} - (\beta_1 + \varepsilon_{t-1} - \varepsilon_{t-2}) = \varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2} \quad 7.13$$

显然，2 阶差分后序列 $\{\nabla^2 x_t\}$ 也是平稳序列，也将原序列 $\{x_t\}$ 中的非平稳趋势充分提取

了。考察它们的方差状况：

$$\text{Var}(\nabla X_t) = \text{Var}(\beta_1 + \varepsilon_t - \varepsilon_{t-1}) = 2\sigma_\varepsilon^2 \quad 7.14$$

$$\text{Var}(\nabla^2 X_t) = \text{Var}(\varepsilon_t - 2\varepsilon_{t-1} + \varepsilon_{t-2}) = 6\sigma_\varepsilon^2 \quad 7.15$$

即，2阶差分后序列 $\{\nabla^2 X_t\}$ 的方差大于1阶差分后序列 $\{\nabla X_t\}$ 的方差，在这个场合下2阶差分就属于过差分，其实质就是因为过多次数的差分导致有效信息的无谓浪费，从而降低了拟合的精度。

三、ARIMA 模型

ARMA 系列拟合模型成立的前提是观测值序列满足弱平稳 Weakly Stationary，即均值 Mean 和方差 Variance 是常数，自协方差 Autocovariance 或自相关系数 Autocorrelation 仅依赖于时间间隔 time lag。但在实际中，大部分的时间序列数据都是非平稳的，即均值或（和）方差是变动的。

1、 $\phi_1 > 1$ 的 AR(1)模型

为了很好的理解 ARIMA 模型，先考察 $\phi_1 > 1$ 的 AR(1)模型：

$$X_t = \phi_1 X_{t-1} + \varepsilon_t \quad 7.16$$

将 X_{t-1} 以式 7.16 的形式代入，得

$$X_t = \phi_1(\phi_1 X_{t-2} + \varepsilon_{t-1}) + \varepsilon_t \quad 7.17$$

再将 X_{t-2} 以式 7.16 的形式展开并代入式 7.17，得到：

$$X_t = \phi_1^2(\phi_1 X_{t-3} + \varepsilon_{t-2}) + \phi_1 \varepsilon_{t-1} + \varepsilon_t \quad 7.18$$

依此类推展开，最终可以得到：

$$X_t = \phi_1^t X_0 + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_1^2 \varepsilon_{t-2} + \cdots + \phi_1^t \varepsilon_0 = \phi_1^t X_0 + \sum_{i=0}^{t-1} \phi_1^i \varepsilon_{t-i} \quad 7.19$$

其中， $\phi_1^t X_0$ 可以视为 $t = 0$ 时刻 X_t 的条件期望或均值，可以看出时间 t 紧密相关。随着时间 $t = 1, 2, \dots$ ，每个时刻 X_t 条件期望的大小都与 $\varepsilon_1, \varepsilon_2, \dots$ 相关，这些条件期望呈现了 X_t 变动趋势，是一种随机的趋势变动（见图 7.2， $\phi_1 = 1.05$ ）。

式 7.19 的方差为：

$$\text{Var}(X_t) = \sigma_\varepsilon^2 \frac{\phi_1^{2t} - 1}{\phi_1^2 - 1} \quad 7.20$$

可以看出，式 7.20 也是时间 t 的函数，随着 $t \rightarrow \infty$ ，方差也趋向于无穷。从以上推导可以看出 $\{X_t\}$ 在均值和方差上都存在着趋势 trend，是非平稳的。

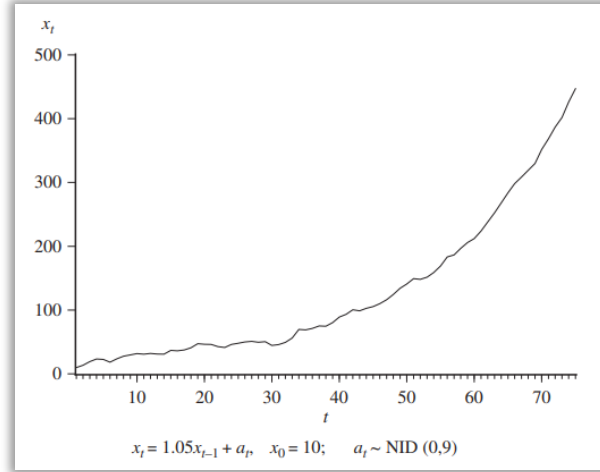


图 7.3 $\phi_1 = 1.05$ 的 AR(1)过程

2、 $\phi_1 = 1$ 的 AR(1)模型 (Random Walk & Random Walk with Drift)

从式 7.16 到式 7.19 可以看出, 当 $\phi_1 > 1$ 时, 其图形是突增性的 (图 7.2), 当 $\phi_1 < 1$ 时, 图形是平稳的 (见第三讲), 现在考察 $\phi_1 = 1$ 的 AR(1)模型 (又称为随机游走模型, Random Walk):

$$X_t = X_{t-1} + \varepsilon_t \quad 7.21$$

再加上截距项, 称为包含漂移项的随机游走模型, Random Walk with Drift, 即:

$$X_t = \theta_0 + X_{t-1} + \varepsilon_t \quad 7.22$$

如果式 7.22 以 $t = 0$ 时刻为起点, 可以转换为:

$$X_t = X_0 + t\theta_0 + \varepsilon_t + \varepsilon_{t-1} + \cdots + \varepsilon_1 + \varepsilon_0 \quad 7.23$$

由此,

$$\mu_t = E(X_t) = X_0 + t\theta_0$$

$$\gamma_{0,t} = \text{Var}(X_t) = t\sigma_\varepsilon^2$$

$$\gamma_{k,t} = \text{Cov}(X_t, X_{t-k}) = E(X_t X_{t-k}) - E(X_t)E(X_{t-k}) = (t-k)\sigma_\varepsilon^2$$

$$\rho_{k,t} = \frac{\gamma_{k,t}}{\sqrt{\gamma_{0,t}\gamma_{0,t-k}}} = \frac{t-k}{\sqrt{t(t-k)}} = \sqrt{\frac{t-k}{t}}$$

如果 t 远大于 k , 则 $\rho_{k,t}$ 接近于 1, 这说明时序 $\{X_t\}$ 的值比较平滑, 同时, μ_t 和 $\gamma_{0,t}$ 会随着 t 而变化, $\{X_t\}$ 是非平稳的。

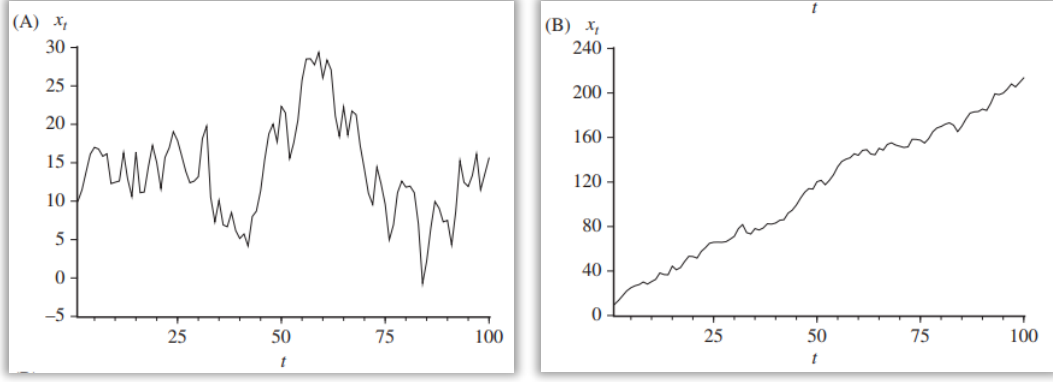


图 7.4 Random Walk (A) 和 Random Walk with Drift (B)

图 7.4 (A) 呈现的是 Random Walk, $X_t = X_{t-1} + \varepsilon_t$, $X_0 = 10$, 图 7.4 (B) 呈现的是 Random Walk with Drift, $X_t = 2 + X_{t-1} + \varepsilon_t$, $X_0 = 10$ 。从模型上看, 二者只相差一个截距项 2, 但图形的差异非常大, 共同的特征是随着时间的推移, X_t 的值和初始值 X_0 之间差异很大。

3、ARIMA 模型

Random Walk 是最简单的非平稳 ARIMA(p, d, q) 模型, 也是最具代表性的模型。式 7.22 通过移项, 可以得到:

$$\nabla X_t = X_t - X_{t-1} = \theta_0 + \varepsilon_t \quad 7.24$$

$\{\varepsilon_t\}$ 序列为白噪声序列, 因此 $\{X_t\}$ 通过 1 阶差分得到的 $\{\nabla X_t\}$ 序列是平稳序列, 推而广之, 非平稳的时序 $\{X_t\}$ 如可以通过 d 阶差分形成新的平稳时间序列 $\{\nabla^d X_t\}$, 而此新平稳时序又采用 ARMA 模型进行构建, 则可以形成针对 $\{X_t\}$ 的模型:

$$\Phi(B)\nabla^d X_t = \Theta(B)\varepsilon_t \quad 7.25$$

。式中, $\nabla^d = (1 - B)^d$, $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$ 为自回归系数多项式, $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$ 为移动平均系数多项式。式 7.25 可以转化为:

$$\nabla^d X_t = \frac{\Theta(B)}{\Phi(B)} \varepsilon_t = \sum_{i=0}^d (-1)^i C_d^i X_{t-i} \quad 7.26$$

说明差分后序列等于原序列的若干序列值的加权和, 且差分平稳序列又可拟合自回归移动平均模型 ARMA 模型, 因此称为求和自回归移动平均模型 Autoregressive Integrated Moving Average Model, ARIMA(p, d, q)。式 7.26 可以看出, ARIMA 模型的实质就是差分运算与 ARMA 模型的组合, 这一关系意义重大, 说明任何非平稳序列如何可以通过适当阶数的差分实现差分平稳, 就可以对差分后序列进行 ARMA 模型拟合。

式 7.25 可以写成：

$$\Phi(B)(1-B)^d X_t = \Theta(B)\varepsilon_t \quad 7.27$$

因为 $\{X_t\}$ 序列 d 阶后平稳，服从 ARMA 模型，则 $1/\lambda_i$ 是系数多项式的根，则：

$$\Phi(B) = (1 - \lambda_1 B)(1 - \lambda_2 B) \dots (1 - \lambda_p B) \quad |\lambda_1| < 1 \quad 7.28$$

令 $\Phi(B)(1-B)^d = 0$ （此处 B 看成是变量），则将式 7.28 代入，得：

$$(1 - \lambda_1 B)(1 - \lambda_2 B) \dots (1 - \lambda_p B)(1 - B)^d = 0 \quad 7.29$$

说明广义自回归系数多项式 $\Phi(B)(1-B)^d$ 共有 $(p+d)$ 个根，其中 p 个根 $1/\lambda_i$, $i = 1, 2, \dots, p$ ，即特征根的倒数，在单元圆外， d 个根在单位圆上。因此，ARIMA(p, d, q)模型不平稳。

4、建模步骤

第一步：模型检验 Model Test。包括平稳性检验和白噪声检验。如果检验是非平稳的，观察时序图特征，进行趋势差分或/和周期差分，再检验差分后时序是否为平稳的非白噪声序列。

第二步：计算差分后序列的样本自相关系数 ACF 和偏自相关系数 PACF；

第三步：模型识别（模型定阶）Model Identification。根据 ACF 和 PACF 的拖尾性和截尾性，选择恰当的 ARMA(p, q)模型对差分后的序列进行拟合。

第四步：参数估计 Parameters Estimation。

第五步：模型检验 Model Test。包括拟合优度检验、参数的有效性检验、模型的有效性检验（残差序列是否为白噪声）。

第六步：模型择优 Model Selection。拟合模型通过检验，仍然可以转向第三步，充分考虑各种可能，建立多个拟合模型，从所有通过检验的拟合模型中选择最优模型。

第七步：模型预测 Model Forecasting。

四、疏系数模型

ARMA(p, q)模型和 ARIMA(p, d, q)模型通常包含 $(p+q)$ 个独立的未知系数， $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$ 。如果该模型中有部分自回归系数 $\phi_j (1 \leq j \leq p)$ 或部分移动平均系数 $\theta_k (1 \leq k \leq q)$ 为零，即 ARMA(p, q)模型和 ARIMA(p, d, q)模型中有部分系数缺省了，那么该模型就称为疏系数模型。教材中 AR(p)、MA(q)、ARMA(p, q)和 ARIMA(p, d, q)模型的定义式均有 $\phi_j \neq 0 (1 \leq j \leq p)$ 和(或) $\theta_k \neq 0 (1 \leq k \leq q)$ ，目的是为了统计性质推导的方便。在实际模型构建

中，疏系数模型时有应用。

疏系数模型的模型识别还是依据自相关系数图和偏自相关系数图的拖尾性和截尾性。具体的操作要看实际运用过程。

五、季节 ARIMA 模型（Seasonal ARIMA 或 SARIMA）

因素分析法告诉我们，时间序列的确定性波动来自于趋势 Trend、季节 Seasonal、循环变动 Cynical 三者。在宏观经济时间序列数据的分析中，循环变动通常是专门用来分析经济周期的，经济周期的时间跨度通常在 5-10 年以上，常规的时间序列分析，很难区分趋势变动和循环变动，由此，将趋势 Trend 和循环变动 Cynical 合二为一，统称为趋势变动。而趋势的分析方法就是趋势差分+ARMA，模型为：

$$\Phi(B)(1-B)^d X_t = \Theta(B)\varepsilon_t$$

如果时序中只包含季节变动信息，分析方法是周期差分+ARMA，模型为：

$$\Phi(B)(1-B^s)X_t = \Theta(B)\varepsilon_t$$

如果时序中既包含趋势信息，又包含季节变动信息，则依据趋势和季节变动的关系，区分为加法模型和乘法模型 Multiplicative SARIMA。加法模型意味着趋势信息和季节变动信息之间是没有关系的，乘法模型意味着趋势信息和季节变动之间是存在着联系的，但事前是无法进行判断的，通常的做法是先采用加法模型，如果残差的白噪声检验无法通过，则用乘法模型。

加法模型如下：

$$(1-B)^d(1-B^s)X_t = \frac{\Theta(B)}{\Phi(B)}\varepsilon_t \quad 7.30$$

简记作 $ARIMA(p, (d, S), q)$ ，或 $ARIMA(p, d, q) \times (0, 1, 0)_s$

乘法模型如下：

$$(1-B)^d(1-B^s)X_t = \frac{\Theta(B) * \Theta_s(B)}{\Phi(B) * \Phi_s(B)}\varepsilon_t \quad 7.31$$

以下举例说明 SARIMA 模型

例 7-1（加法模型），拟合 1962 年~1991 年德国工人季度失业率序列。

第一步：按照所给数据，绘制观察值序列的时序图，见图 7.1。

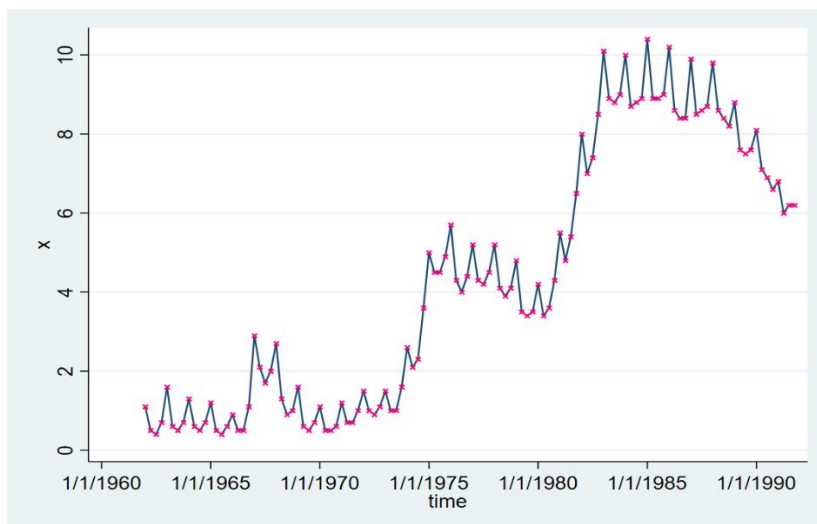


图 7.1 德国工人季度失业率序列时序图

图 7.1 显示，该序列既含有长期趋势又含有以年为周期的季节效应，其中 1 季=3 个月。

第二步：差分平稳。对原序列做 1 阶差分消除趋势，再做 4 步差分消除季节效应。差分后的时序图，见图 7.2。

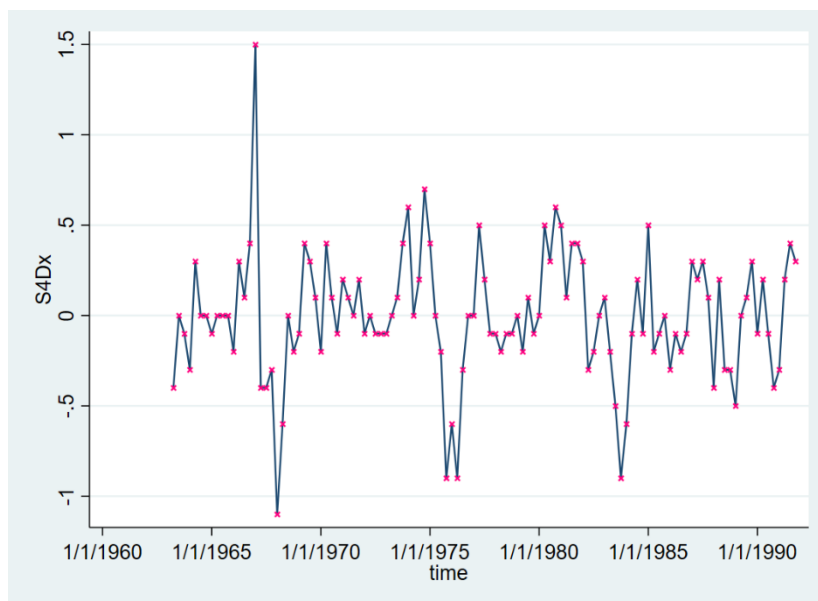


图 7.2 德国工人季度失业率 1 阶 4 步差分后序列时序图

图 7.2 显示，差分后的序列已无明显趋势或周期，随机波动比较平稳。

第三步：对差分后的序列进行单位根检验和白噪声检验。

ADF 检验，采用的是类型一，即不包含常数项的回归结构，延迟阶数为 0。计算得到 t 统计量值为 -6.773，对应的 1% 的临界值 Critical Value 为 -2.598，5% 的临界值 Critical Value

为-1.950，10%的临界值 Critical Value 为-1.611。 τ 统计量值远小于1%的临界值，说明原假设 H_0 : 存在单位根（即序列非平稳）这个事件发生的概率远小于0.01。结论：差分后的序列是平稳的。白噪声检验采用 LB 统计量：Portmanteau (Q) statistic = 43.8369, Prob > chi2(6) = 0.0000，拒绝原假设 $H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0$ ，结论：非白噪声。

第四步：模型定阶。观察 ACF 图和 PACF 图

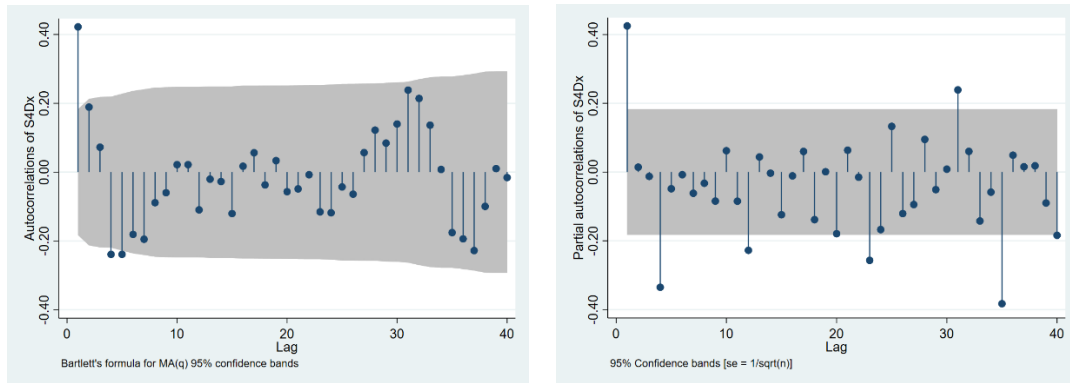


图 7.3 德国工人季度失业率 1 阶 4 步差分后道的 ACF 图（左）和 PACF 图（右）

考察图 7.3 的图形特征，ACF 显示出明显的下滑轨迹，类似“坐滑梯”，这是典型的拖尾特征。PACF 除了 1 阶和 4 阶明显大于 2 倍标准差，其他阶数的值基本上都在 2 倍标准差范围内波动，所以可以尝试拟合疏系数模型 AR(1,4)，当然这是对 1 阶 4 步差分后的序列。对原序列应该是拟合疏系数季节加法模型 ARIMA((1,4),(1,4),0)。

第五步：参数估计。这里采用 MLE 估计法，得到的模型为：

$$(1-B)(1-B^4)x_t = \frac{1}{1-0.445B+0.272B^4} \varepsilon_t \quad \text{Var}(\varepsilon_t) = 0.304^2 \quad 7.32$$

第六步：模型检验。分别是参数的有效性检验合模型的有效性检验。

计算得到参数有效性检验的 t 统计量的值分别为 8.68 和 -3.97，其绝对值远大于 1.96（95% 概率保证程度下的 t 值），因此，拒绝原假设，即参数显著非零。

计算得到残差白噪声检验的 Portmanteau (Q) statistic = 2.2087, Prob > chi2(6) = 0.8995，接受原假设 $H_0: \rho_1 = \rho_2 = \dots = \rho_k = 0$ ，结论：残差序列是白噪声，说明式 7.32 为德国工人季度失业率时序的有效模型。

例 7-2（乘法模型），拟合 1948-1981 年美国女性（20 周岁以上）月度失业率序列。

第一步：按照所给数据，绘制观察值序列的时序图，见图 7.4。

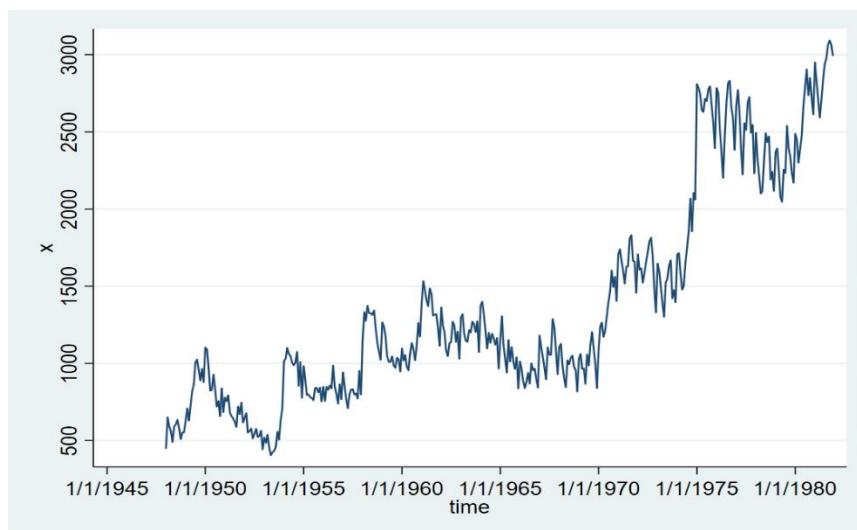


图 7.4 美国女性月度失业率序列时序图

图 7.4 显示，该序列既含有长期递增趋势又含有以年为周期的季节效应，其中 1 季=1 个月。

第二步：差分平稳。对原序列做 1 阶差分消除趋势，再做 12 步差分消除季节效应。差分后的时序图，见图 7.5。

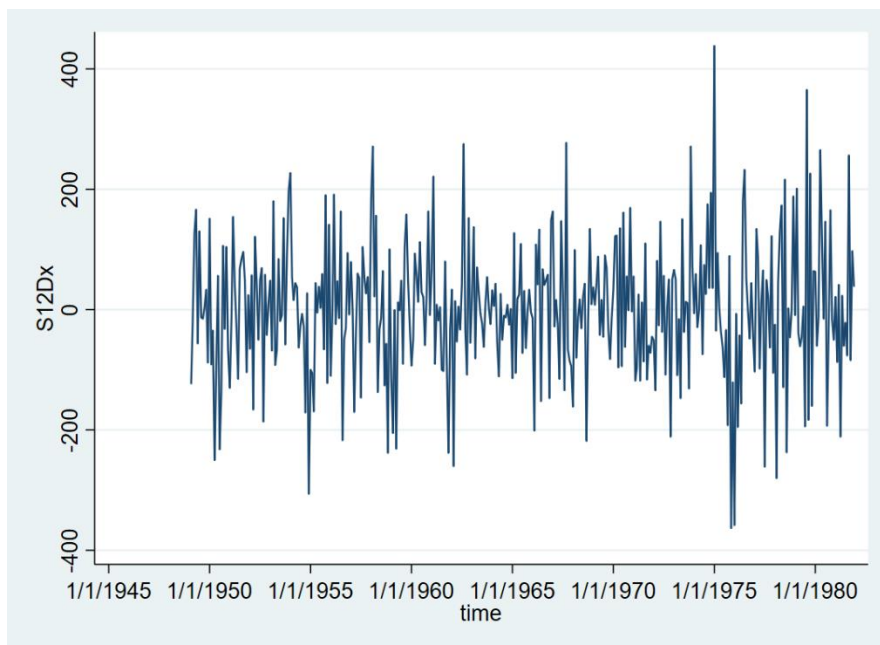


图 7.5 美国女性月度失业率序列时序图 1 阶 12 步差分后序列时序图

第三步：对差分后的序列进行单位根检验和白噪声检验。

这里具体的 ADF 检验和 Q 检验的结果略过，结果：平稳非白噪声。

第四步：模型定阶。观察 ACF 图和 PACF 图（图 7.6）。

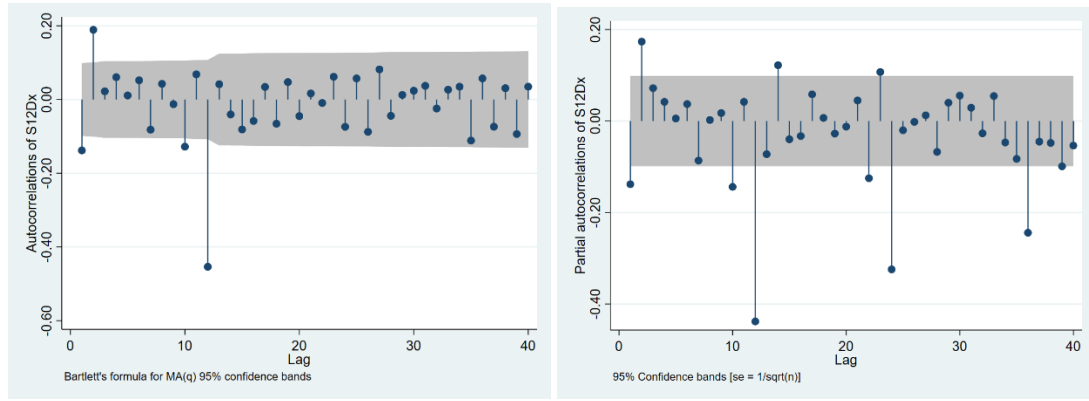


图 7.6 德国工人季度失业率 1 阶 4 步差分后道的 ACF 图（左）和 PACF 图（右）

ACF 图和 PACF 图显示延迟 12 阶自相关系数显著大于 2 倍标准差，这说明差分后序列中仍然蕴涵非常明显的季节效应。延迟 1 阶、2 阶的自相关系数也大于 2 倍标准差，这说明序列还具有短期相关性。依据图 7.6 的性质，尝试拟合 $AR(1,12)$ ， $MA(1,2,12)$ 和 $ARMA((1,12)(1,12))$ ，但拟合的残差均无法通过白噪声检验。反复尝试的结果均不理想，说明简单的季节加法模型并不适合拟合这个序列，可以认为该序列的季节效应和短期效应之间具有复杂的关联性，这里假定短期相关性和季节效应之间具乘法关系。

乘法模型的构造原理如下：序列具有短期相关时，采用低阶的 $ARMA(p,q)$ 模型提取，序列具季节效应，且季节效应本身还具相关性时，可以采用以周期步长 S 步为单位的 $ARMA(P,Q)_S$ 模型提取。由于短期相关性和季节效应之间具有乘积关系，所以拟合的模型应该是 $ARMA(p,q)$ 和 $ARMA(P,Q)_S$ 的乘积。综合之前 d 阶差分 and S 步差分，原序列拟合的乘法模型完整结构如下：

$$(1-B)^d(1-B^S)X_t = \frac{\Theta(B) * \Theta_S(B)}{\Phi(B) * \Phi_S(B)} \varepsilon_t$$

具体到图 7.6 的定阶，首先考虑 1 阶 12 步差分之后序列 12 阶以内的自相关系数和偏自相关系数的特征，以确定短期相关模型。ACF 图和 PACF 图显示 12 阶以内的自相关系数和偏自相关系数均不截尾，所以尝试使用 $ARMA(1,1)$ 模型提取差分后序列的短期自相关信息。再考虑季节自相关特征，这时考察延迟 12 阶、24 阶等以周期长度为单位的 ACF 和 PACF 的特征。ACF 图显示延迟 12 阶自相关系数显著非零，但是延迟 24 阶自相关系数落入 2 倍标准差范围。而 PACF 图显示延迟 12 阶和延迟 24 阶的偏自相关系数都显著非零。所以，可以认为季节自相关特征是 ACF 截尾，PACF 拖尾，这时用以 12 步为周期的 $MA(1)$ 模型提取差分后序列的季节相关信息。综上可得模型的具体结构为：

$$(1-B)(1-B^{12})X_t = \frac{(1-\theta_1)}{(1-\phi_1)}(1-\theta_{12}B^{12})\varepsilon_t \quad 7.33$$

第五步：参数估计。这里采用 MLE 估计法，得到的模型为：

$$(1-B)(1-B^{12})X_t = \frac{(1-0.604)}{(1+0.727)}(1+0.792B^{12})\varepsilon_t \quad Var(\varepsilon_t) = 86.28^2 \quad 7.34$$

第六步：模型检验。分别是参数的有效性检验合模型的有效性检验，即对参数进行 t 检验以及对残差序列进行白噪声检验，这里都通过检验。具体过程略。