第 六 章 季节模型



学习目标与要求

- ■了解季节模型的解构和特征.
- 理解乘积季节自回归移动平均模型的解构和特征.
- ■理解乘积季节求和自回归移动平均模型的解构和特.
- 掌握几类特殊的季节模型的建模过程.



本章结构

- 1. 简单季节自回归移动平均模型
 - 2. 乘积季节自回归移动平均模型
- 3. 季节求和自回归移动平均模型
- 4. 季节求和自回归移动平均模型的预测



• 简单季节自回归移动平均模型

在应用时间序列的很多领域中,特别是许多商业和经济领域中的数据都呈现出每隔一段时间重复、循环的季节现象. 在前面,我们曾用确定性趋势方法对这类季节性数据建模,但是这样建立起来的确定性季节模型不能充分解释此类时间序列的行为,而且模型残差仍然在许多滞后点高度自相关.从本节开始介绍随机季节模型,可以看到该模型可以很好地拟合季节性数据,有效地克服确定性季节模型的不足.首先讨论平稳季节模型,之后再研究非平稳的情况.

- 季节移动平均模型
- 一般地,用 s 表示季节周期: s = 12表示月度数据, s = 4表示季度数据.

非平稳序列的定义

考虑由下式生成的时间序列:

$$y_{t} = \varepsilon_{t} + \theta \varepsilon_{t-12}$$
, 注意到

$$Cov(y_t, y_{t-i}) = Cov(\varepsilon_t + \theta \varepsilon_{t-12}, \varepsilon_{t-i} + \theta \varepsilon_{t-12-i}) = 0, \quad i = 1, 2, \dots, 11.$$

$$\operatorname{Cov}(y_{t}, y_{t-12}) = \operatorname{Cov}(\varepsilon_{t} + \theta \varepsilon_{t-12}, \varepsilon_{t-12} + \theta \varepsilon_{t-24}) = \theta \sigma_{\varepsilon}^{2}.$$

易见, 序列 $\{y_t\}$ 平稳且仅在延迟 12 期才具有非零自相关函数.

概括上述想法, 我们称具有如下结构的模型为季节周期为 s 的 q 阶季节移动平均模型, 简记为 MA(q) s ,

$$y_{t} = \varepsilon_{t} - \theta_{1} \varepsilon_{t-s} - \theta_{2} \varepsilon_{t-2s} - \dots - \theta_{q} \varepsilon_{t-qs}$$

其季节移动平均 (MA) 系数多项式为

$$\Theta(B) = 1 - \theta_1 B^s - \theta_2 B^{2s} - \dots - \theta_q B^{qs}.$$



若 $\Theta(B)$ =0的根都在单位圆外,则该模型可逆.显然,该序列总是平稳的,且其自相关函数只在 s,2s,3s,・・・,qs 等季节滞后点上非零. 特别地

$$\rho k_{s} = \frac{\theta_{k} + \theta_{1}\theta_{k+1} + \theta_{2}\theta_{k+2} + \dots + \theta_{q-k}\theta_{q}}{1 + \theta_{1}^{2} + \theta_{2}^{2} + \dots + \theta_{q}^{2}}, \quad k = 1, 2, \dots, q.$$

当然,季节 MA(q) s 模型可以看做阶数 qs 的非季节性 MA模型的特例,即后者除了在季节滞后点 s,2s,3s,···,qs 处非零外,其余所有的 θ 值都取零.

■ 季节自回归模型

类似地可定义季节自回归模型的概念. 首先考察一个特例. 考虑 $y_t = \phi y_{t-12} + \varepsilon_t$ (6.1)

其中, $| \varphi | < 1$, 且 ε t 与 $y_{t-i}(i=1,2,3,\cdots)$ 独立. 容易证明, $| \varphi | < 1$ 保证了平稳性, 而且显然 $E(y_t)=0$. (6.1)式两边同时乘以 y_{t-k} , 取期望后, 再除以 $\gamma(0)$ 得

$$\rho(k) = \phi \rho(k-12), k \ge 1.$$
 (6.2)

由 (6.2) 式得到 $\rho(12) = \phi \rho(0) = \phi, \rho(24) = \phi \rho(12) = \phi^2, \dots, \rho(12k) = \phi^k, \dots$

在 (6.2) 式中, 分别令 k=1 和k=11 ,由 $\rho(k)=\rho(-k)$ 得 $\rho(1)=\phi\rho(11), \rho(11)=\phi\rho(1).$

从而 $\rho(1) = \rho(11) = 0$. 类似可证, 除了在季节滞后12,24,36,… 处以外, $\rho(k)$ 全为零. 而滞后 12,24,36,… 期的自相关函数表现出类似 AR(1) 模型的指数衰减.

一般地, 我们称具有如下结构的模型为季节周期为 s 的 p 阶



季节自回归模型,简记为AR(P)s,

$$y_{t} = \phi_{1} y_{t-s} + \phi_{2} y_{t-2s} + \dots + \phi_{p} y_{t-ps} + \varepsilon_{t},$$
 (6.3)

其中,季节自回归 (AR) 系数多项式为

$$\Phi(B) = 1 - \phi_1 B^s - \phi_2 B^{2s} - \dots - \phi_p B^{ps}.$$

若 $\Phi(B) = 0$ 全部根的绝对值均大于 1,则该模型为平稳的. 模型 (6.3) 可以看做是一个阶数 的特定 AR(p) 模型,仅在季节滞后 s,2s,・・・, 处才有非零的 ϕ 系数.



本章结构

- 1. 简单季节自回归移动平均模型
 - 2. 乘积季节自回归移动平均模型
- 3. 季节求和自回归移动平均模型
- 4. 季节求和自回归移动平均模型的预测



■ 乘积季节自回归移动平均模型

将季节模型和非季节模型相结合,可以构造出既包括季节延迟自相关又包括低阶临近延迟自相关的简约模型 —— 乘积季节模型.

考虑一个 MA 模型, 其系数多项式如下:

$$(1+\theta B)(1+\Theta B^{12})=1+\theta B+\Theta B^{12}+\theta \Theta B^{13},$$

相应的时间序列模型为

$$y_{t} = \varepsilon_{t} + \theta \varepsilon_{t-1} + \Theta \varepsilon_{t-12} + \theta \Theta \varepsilon_{t-13}$$



考虑一个 MA 模型, 其系数多项式如下:

$$(1+\theta B)(1+\Theta B^{12})=1+\theta B+\Theta B^{12}+\theta\Theta B^{13},$$
相应的时间序列模型为

$$y_{t} = \varepsilon_{t} + \theta \varepsilon_{t-1} + \Theta \varepsilon_{t-12} + \theta \Theta \varepsilon_{t-13}$$

可验证此模型的自相关函数仅在延迟后 1, 11, 12, 13 期非零, 即 $\gamma(0) = (1+\theta^2)(1+\Theta^2)\sigma_{\varepsilon}^2$,

$$\rho(1) = \frac{\theta}{1+\theta^2}, \rho(12) = \frac{\Theta}{1+\Theta^2}, \rho(11) = \rho(13) = \frac{\theta\Theta}{(1+\theta^2)(1+\Theta^2)}.$$

将这种想法一般化,可以构造一类乘积季节模型.一般地,称 具有如下结构的模型为季节周期为 s 的乘积季节 模型: $ARMA(p,q)\times(P,Q)_s$

$$\phi(B)\Phi(B)x_{t} = \theta(B)\Theta(B)\varepsilon_{t}, \varepsilon_{t} \sim WN(0,\sigma^{2})$$

其中, AR 系数多项式和 MA 系数多项式分别为 $\phi(B)$ $\Phi(B)$,这里

$$\phi(B) = 1 - \phi_{1}B - \phi_{2}B^{2} - \dots - \phi_{p}B^{p},$$

$$\Phi(B) = 1 - \Phi_{1}B^{s} - \Phi_{2}B^{2s} - \dots - \Phi_{p}B^{ps},$$

$$\theta(B) = 1 - \theta_{1}B - \theta_{2}B^{2} - \dots - \theta_{q}B^{q},$$

$$\Theta(B) = 1 - \Theta_{1}B^{s} - \Theta_{2}B^{2s} - \dots - \Theta_{0}B^{Qs}.$$

乘积季节ARMA(p,q)×(P,Q)_s模型,可以看做 AR 阶数为 q+Q_s的ARMA 模型的特例,此模型仅有 p+P+q+Q 个系数非零.

特别地, 当 P = q = 1, p = Q = 0, s = 12时, 我们得到 $ARMA(0,1)\times(1,0)_{12}$ 模型: $x_t = \Phi x_{t-12} + \varepsilon_t - \theta \varepsilon_{t-1}$. 经过计算得到



$$\gamma(1) = \Phi \gamma(11) - \theta \sigma_{\varepsilon}^{2},$$
$$\gamma(k) = \Phi \gamma(k-12), k \ge 2.$$

进而可得

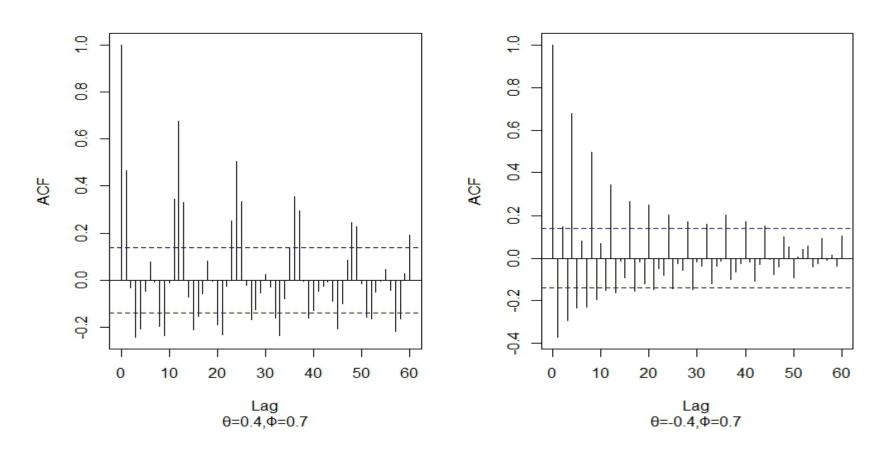
Ħ.

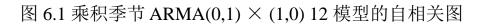
$$\gamma(0) = \left(\frac{1+\theta^2}{1-\Phi^2}\right)\sigma_{\varepsilon}^2, \rho(12k-1) = \rho(12k+1) = \left(-\frac{\theta}{1-\theta^2}\Phi^k\right), k \ge 0.$$

当延迟阶数为其他值时,自相关函数为零.

图 6.1 为 θ = ±0.4, Φ = 0.7 时, 乘积季节ARMA(0,1)×(1,0)₁₂模型的自相关函数图. 对比两图可以发现, 包含自回归成份的乘积季节模型 ARMA(0,1)×(1,0)₁₂模型的自相关函数并不会出现截尾特征, 而是在延迟 12k-1,12k和 12k+1(k=1,2,···) 快速衰减. 这是许多季节时间序列的样本自相关函数图的典型特征.









本章结构

- 1. 简单季节自回归移动平均模型
- 2. 乘积季节自回归移动平均模型
- 3. 季节求和自回归移动平均模型
- 4. 季节求和自回归移动平均模型的预测



■ 乘积季节求和自回归移动平均模型 季节差分是非平稳季节序列建模的一个重要工具。 设 s 是一个季节时间序例 } 的季节周期,则它的 周期为 s 的季节差分为 = x_t - x_{t-s}.

值得注意的是,长度为n的序列其季节差分序列的长度为s,即季节差分后丢失 s个数据值.

对于一些非平稳的时间序列来讲, 经过 d 阶差分和 D 阶季节差分后, 可以变成平稳的时间

序列 $\{y_t\}$ 即 $y_t = \nabla^d \nabla^D_s x_t$.

若 $\{y_t\}$ 满足季节周期为s的ARMA $(p,q)\times(P,Q)_s$ 模型,



那么称 $\{x t\}$ 为季节周期为 s、非季节阶数为 p,d,q、季节阶数为 P,D,Q 的乘积季节求和自回归移动平均模型,记作 SARIMA(p,d,q)× $(P,D,Q)_s$,即 $\{x t\}$ 满足

$$\phi(B)\Phi(B)\nabla^{d}\nabla_{s}^{D}X_{t}=c+\theta(B)\Theta(B)\varepsilon_{t},\varepsilon_{t}\sim WN(0,\sigma^{2}),$$

其中, c 为常数, $\phi(B)$ 和 $\theta(B)$ 分别为自回归系数多项式和移动平均系数多项式; $\Phi(B)$ 和 $\Theta(B)$ 分别为季节自回归系数多项式和季节移动平均系数多项式; $\nabla^d = (1-B)^d$ 且 $\nabla^D_s = (1-B^s)^D$.

上述模型包含了一族范围很广,而且表示灵活的模型类.实践表明,这些模型能够充分拟合许多序列.

乘积季节求和自回归移动平均模型的建模季节模型的识别、拟合和检验方法同以前介绍过的方法基

本一致. 下面通过举例说明季节模型的建模过程.

例 6.1 对 1966 年 1 月至 1990 年 12 月夏威夷 CO_2 排放量数据进行分析, 并建立拟合模型.

解下面我们按照建模步骤来分析 1966 年 1 月至 1990 年 12 月夏威夷 *CO*₂ 排放量数据.

1. 模型识别

首先读入数据,并绘制时序图.根据图 6.2 中的时序图可知,1966 年 1 月至 1990 年 12 月夏威夷 CO_2 排放量呈现季节性递增趋势.然后,对该排放量观察值序列作 1 阶差分以消除线性递增趋势.具体命令如下,运行结果见图 6.2.



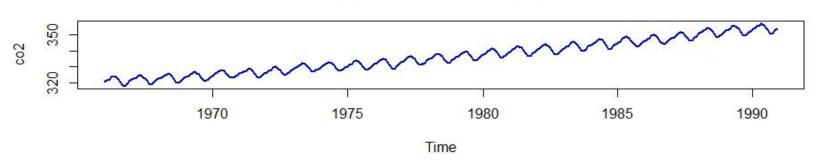
```
> co2 <- read.table(file="E:/DATA/CHAP6/1.csv",sep=",",header=T)
> co2 <- ts(co2$Carbondioxide,start=c(1966,1),frequency = 12)
> par(mfrow=c(2,1))
> plot(co2,lwd=2,col="blue",main="夏威夷二氧化碳排放量时序图")
> plot(diff(co2,1,1),lwd=2,main="夏威夷二氧化碳排放量一阶差分时序图")
```

由图 6.2 中的 1 阶差分时序图可见, 经过 1 阶差分运算, 该序列已无递增趋势, 但是有明显的季节性周期. 故继续对 1 阶差分序列作周期为 12 的季节差分, 并且绘制时序图. 具体命令如下, 运行结果见图 6.3.

```
> par(mfrow=c(1,1))
> x <- diff(diff(co2),12,1)
> plot(x)
```



夏威夷二氧化碳排放量时序图



夏威夷二氧化碳排放量 1 阶差分时序图

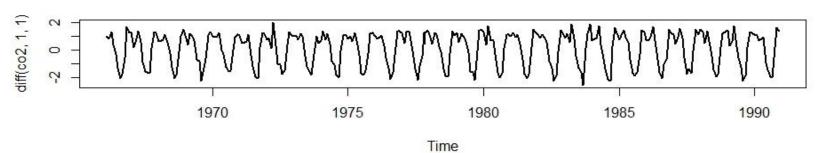


图 6.2 1966 年 1 月至 1990 年 12 月夏威夷 CO 2 排放量图



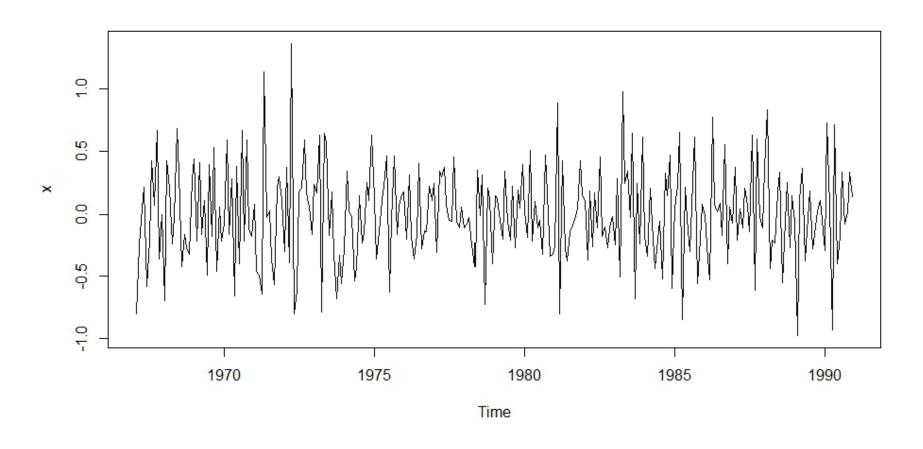


图 6.3 夏威夷二氧化碳排放量季节差分序列时序图

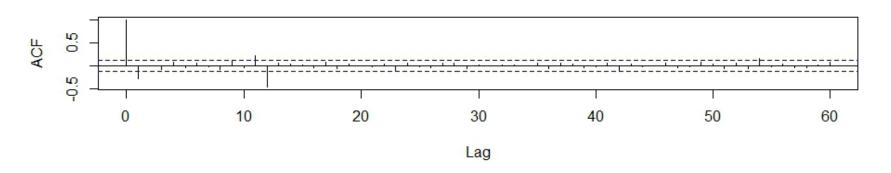


由图 6.3 中的季节差分时序图可见,序列季节差分之后已经没有显著的季节性了,并且表现出一定平稳序列的特征.为了定阶,作出差分序列的自相关函数图和偏自相关函数图.具体命令如下,运行结果如图 6.4 所示.

- > par(mfrow=c(2,1))
- > acf(as.vector(x),lag=60,main="夏威夷二氧化碳排放量差分序列自相关函数图")
- > pacf(as.vector(x),lag=60,main="夏威夷二氧化碳排放量差分序列偏自相关函数图")



夏威夷二氧化碳排放量差分序列自相关函数图



夏威夷二氧化碳排放量差分序列偏自相关函数图

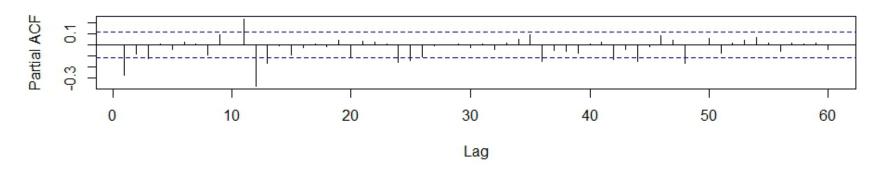


图 6.4 夏威夷二氧化碳排放量差分序列相关图



从图 6.4 可见, 自相关函数具有延迟 12 期的季节周期, 且在季节点上的自相关函数具有 1阶截尾性, 季节内的自相关函数也具有截尾性. 同时, 偏自相关函数无论是季节点上还是季节内各延迟点上都显示了显著的拖尾性. 因此选择乘积季节模型 SARIMA(0,1,1) × (0,1,1) 来拟合数据.

2. 模型拟合

在 R 语言中,可以用函数 sarima() 来拟合季节模型. 在调用函数 sarima() 之前,需要安装并加载程序包 astsa. 函数 sarima() 的命令格式如下:

sarima(x, p, d, q, P =, D =, Q =, S =, details =)

该函数的参数说明:



- x: 拟合变量名.
- p: 自回归阶数.
- d: 差分阶数.
- q: 移动平均阶数.
- P: 季节自回归阶数, 仅用于季节模型.
- D: 季节差分阶数, 仅用于季节模型.
- Q: 季节移动平均阶数, 仅用于季节模型.
- S: 季节周期数, 仅用于季节模型.
- details: 非线性优化结果输出按钮. details=TRUE 表示输出结果, 是默认值; details=FALSE 表示不输出结果.



用函数 sarima() 来拟合 1966 年 1 月至 1990 年 12 月 夏威夷 CO 2 排放量数据. 具体命令及运行结果如下:

```
> co2.fix <- sarima(co2,0,1,1,P=0,D=1,Q=1,S=12,details=F)
> co2.fix
$fit
```

Call:

```
stats::arima(x = xdata, order = c(p, d, q), seasonal = list(order = c(P, D, Q), period = S), include.mean = !no.constant, optim .control = list(trace = trc, REPORT = 1, reltol = tol))
```

Coefficients:

```
ma1 sma1
-0.3369 -0.8614
s.e. 0.0624 0.0465
```

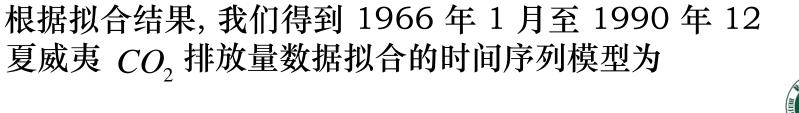
sigma^2 estimated as 0.07766: log likelihood = -48.72, aic = 103.45



```
$degrees_of_freedom
[1] 298
$ttable
        Estimate SE t.value p.value
ma1 -0.3369 0.0624 -5.4022
sma1 -0.8614 0.0465 -18.5082
                                  0
$AIC
[1] -1.54205
$AICc
[1] -1.535113
$BIC
```

月夏威夷 CO, 排放量数据拟合的时间序列模型为

[1] -2.517358



$$(1-B^{12})(1-B)x_t = (1+0.8614B^{12})(1+0.3369B)\varepsilon_t, \varepsilon_t \sim N(0,0.07766). \tag{6.4}$$

也可看出,移动平均系数和季节移动平均系数都显著异于零.

3. 诊断检验

从图 6.5 中可以看到,标准化残差图、残差自相关图、QQ 图和 Ljung-Box 统计量图都显示残差不存在显著相关性,即残差为白噪声序列. 因此拟合方程 (6.10) 通过诊断检验,拟合效果良好.



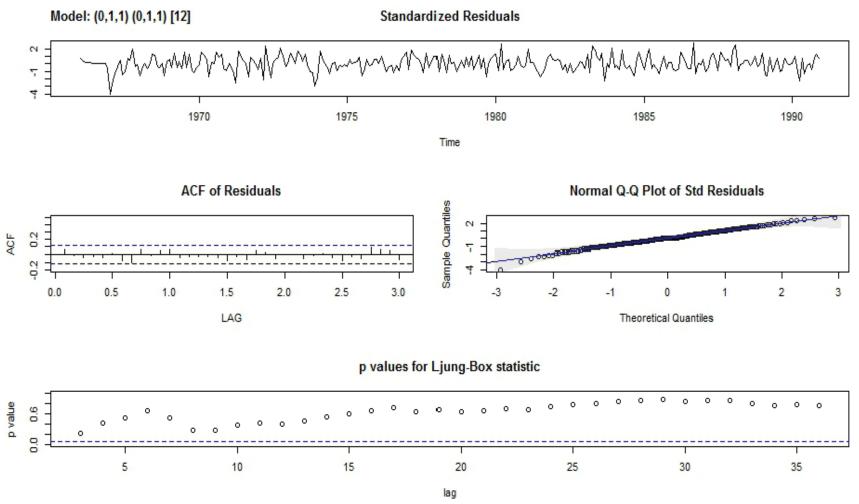


图 6.5 残差诊断图



本章结构

- 1. 简单季节自回归移动平均模型
- 2. 乘积季节自回归移动平均模型
- 3. 季节求和自回归移动平均模型
- 4. 季节求和自回归移动平均模型的预测



季节 ARIMA 模型预测的最简单方法是对模型递归地应用差分方程形式. 下面举例说明.

1. 季节 AR(1) 12 模型的预测

季节 AR(1) 12 模型为

$$x_{t} = \Phi x_{t-12} + \varepsilon_{t},$$

于是可以得到向前1步的预测值为

$$\hat{x}_{t+l} = \Phi \hat{x}_{t+l-12}.$$

利用上式,向前迭代1步得

$$\hat{x}_{t+l} = \Phi^{k+1} x_{t+l-12(k+1)}.$$

如果最后一个观测值在当年 12 月份,那么下一年 1 月份的 预测值为 Φ 乘以当年 1 月份的观测值.下一年 2 月份的预测

值为 Φ 乘以当年 2 月份的观测值. 以此类推, 下一年 1 月份的预测值为 Φ^2 乘以前年 1 月份的观测值.

如果 | Φ | < 1, 那么未来预测值以指数方式衰减. 根据 (5.14) 式可知, 预测方差可以写成

$$\operatorname{Var}\left[e_{t}(l)\right] = \frac{1 - \Phi^{2k+2}}{1 - \Phi^{2}} \sigma_{\varepsilon}^{2},$$

其中 $k \, \mathbb{E}(l-1)/12$ 的整数部分.

2. 季节 MA(1) 12 模型的预测

季节 MA(1) 12 模型为 $x_t = \mu + \varepsilon_t + \Theta \varepsilon_{t-12}$.

从而得其预测值为

$$\hat{x}_{t+1} = \mu + \Theta \varepsilon_{t-11, \hat{x}_{t+2}} = \mu + \Theta \varepsilon_{t-10, \dots, \hat{x}_{t+12}} = \mu + \Theta \varepsilon_t$$



和
$$\hat{x}_{t+l} = \mu, l > 12$$

可见,该模型能够给出第一年中各个月份的不同预测值,然 而一年之后的所有的预测值都由序列均值给出.同样由(5.14) 式给出预测方差

$$\operatorname{Var}\left[e_{t}\left(l\right)\right] = \begin{cases} \sigma_{\varepsilon}^{2}, & 1 \leq l \leq 12, \\ \left(1 + \Theta^{2}\right)\sigma_{\varepsilon}^{2}, & l > 12 \end{cases}$$

3.SARIMA(0,1,1)×(1,0,1)₁₂模型的预测

SARIMA(0,1,1)×(1,0,1)₁₂ 模型:

$$x_{t} - x_{t-1} = \phi(x_{t-12} - x_{t-13}) + \varepsilon_{t} - \theta\varepsilon_{t-1} - \Theta\varepsilon_{t-12} + \theta\Theta\varepsilon_{t-13}.$$

由上式可得 $x_t = x_{t-1} + \phi x_{t-12} - \phi x_{t-13} + \varepsilon_t - \theta \varepsilon_{t-1} - \Theta \varepsilon_{t-12} + \theta \Theta \varepsilon_{t-13}$.



于是当前时刻t的下一步预测值为

$$\hat{x}_{t+1} = x_t + \phi x_{t-11} - \phi x_{t-12} - \theta \varepsilon_t - \Theta \varepsilon_{t-11} + \theta \Theta \varepsilon_{t-12}.$$

再下一步的预测值为

$$\hat{\boldsymbol{x}}_{t+2} = \hat{\boldsymbol{x}}_{t+1} + \phi \boldsymbol{x}_{t-10} - \phi \boldsymbol{x}_{t-11} - \Theta \boldsymbol{\varepsilon}_{t-10} + \theta \Theta \boldsymbol{\varepsilon}_{t-11}.$$

以此类推, 向前预测 $l, l = 1, 2, \dots, 13$, 步时, 噪声项 $\varepsilon_{t-13}, \varepsilon_{t-12}, \dots, \varepsilon_{t}$

进入预测表达式, 但是对于 l=13, 有回归公式

$$\hat{x}_{t+l} = \hat{x}_{t+l-1} + \phi \hat{x}_{t+l-12} - \phi \hat{x}_{t+l-13}, \quad l > 13.$$

4.SARIMA(0,0,0)×(0,1,1), 模型的预测

SARIMA(0,0,0)×(0,1,1)₁₂模型为
$$x_t - x_{t-12} = \varepsilon_t - \Theta \varepsilon_{t-12}$$
.

或者将上式写成
$$x_{t+l} = x_{t+l-12} + \varepsilon_{t+l} - \Theta\varepsilon_{t+l-12}$$
.



据此,我们可以得到所有1月份的预测,同样可以得到所有2月份的预测,等等.

在R语言中,可以用函数 sarima.for()来预测季节模型. 在调用函数 sarima.for()之前,需要安装并加载程序包 astsa. 函数 sarima.for()的命令格式以及主要参数使用方法与函数 sarima()几乎一样,故略去.

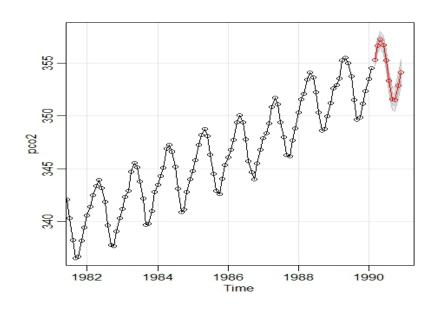
例 6.2 选择模型 SARIMA(0,1,1)×(1,0,1)₁₂ 拟合 1966 年 1 月至 1990 年 2 月夏威夷 CO_2 排放量数据. 预测 1990 年 3 月至 1990 年 12 月的排放量, 并与实际值对比.

解 读入数据, 并且截取 1966 年 1 月至 1990 年 2 月的排放量数据, 应用截取的数据拟合模型 $SARIMA(0,1,1) \times (1,0,1)_{12}$ 具体命令如下, 运行结果如图 6.6 所示.

- > co2 <- read.table(file="E:/DATA/CHAP6/1.csv",sep=",",header=T)</pre>
- > co2 <- ts(co2\$Carbondioxide, start=c(1966, 1), frequency = 12)
- > pco2 <- window(co2,start=c(1966,1),end=c(1990,2)) #截取数据
- >co2.fore <- sarima.for(pco2,10,0,1,1,P=0,D=1,Q=1,S=12)#作10期预测
- 将二氧化碳排放量的预测值与实际值进行比较. 具体命令如下, 运行结果见图 6.7.
- > EXP <- c("实际值","预测值")
- > ts.plot(co2,co2.fore\$pred,gpars=list(lty=c(1,2)),col=c("red",



- + "blue"),lwd=c(1,3))
- > legend(x="topleft",EXP,lty=c(1,2))



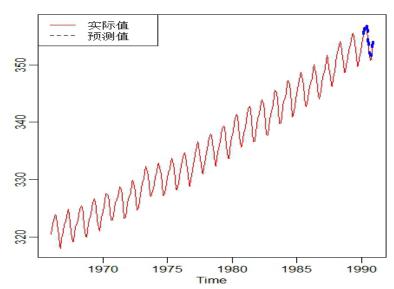


图 6.6 二氧化碳排放量预测图

图 6.7 预测值与实际值比较图

