**38. 非平稳时间序列的确定性分析**

实际中大多数时间序列是非平稳的，对非平稳时间序列的分析方法主要有两类：确定性分析和随机性分析。

确定性分析——提取非平稳时间序列明显的规律性（长期趋势、季节性变化、周期性），目的是：①克服其它因素影响，单纯测度出单一确定因素对序列的影响；②推断各种确定性因素彼此之间相互作用关系及它们对序列的综合影响。

随机性分析——分析非平稳时间序列由随机因素导致的随机波动性。

**（一）趋势分析**

有的时间序列具有明显的长期趋势，趋势分析就是要找出并利用这种趋势对序列发展做出合理预测。

1. 趋势拟合法

即把时间作为自变量，相应的序列观察值作为因变量，建立序列值随时间变化的回归模型。分为线性拟合和非线性拟合。

2. 平滑法

利用修匀技术，消弱短期随机波动对序列的影响，使序列平滑化，从而显示出长期趋势变化的规律。

（1）移动平均、加权移动平均

已知序列值*x*1, …, *xt-*1, 预测*xt*的值为



称为*n*期移动平均值，*n*的选取带有一定的经验性，*n*过长或过短，各有利弊，也可以根据均方误差来选取。

一般最新数据更能反映序列变化的趋势。因此，要突出新数据的作用，可采用加权移动平均法：



其中，.

（2）二次移动平均

对应线性趋势，移动平均拟合值有滞后性，可以采用二次移动平均加以改进：对移动平均值再做一次移动平均。

（3）指数平滑法

指数平滑法是一种对过去观察值加权平均的特殊形式，观测值时间越远，其权数呈指数下降。一次指数平滑法可用于对时间序列进行修匀，以消除随机波动。预测公式为：



其中α∈(0, 1)为平滑常数，为第t期平滑预测值，初始预测值（通常取最初几个实测数据的均值）。

一般来说，时间序列有较大的随机波动时，宜选择较大的α值，以便能较快跟上近期的变化；也可以利用预测误差选择。

（4）二次、三次指数平滑法

即对一次指数平滑后的序列再做一次指数平滑，但不是直接将二次指数平滑值作为预测值，而是利用其来求出方程参数，利用滞后偏差的规律来建立直线趋势模型。计算公式：





, 



其中，m为预测超前期数，取.

（5）霍尔特双参数线性指数平滑法

设α, β∈(0, 1)为参数，为趋势增量。用趋势增量来修正，消除了滞后性，对数据进行平滑：



用指数平滑法估计趋势增量，对相邻两次平滑之差做修正，再加上前期趋势增量，对趋势进行平滑：



计算超前m期的预测值：



初值的选取：, .

**（二）时间序列的分解**

**一、Gramer分解定理**

1963年，Gramer在Wald分解定理的基础上，得到了Gramer分解定理：

任一时间序列{Xt}都可以分解为叠加的两部分：由多项式决定的确定性趋势成分，平稳的零均值误差成分，即



其中，为0均值白噪声序列，B为延迟算子，且





即均值序列反映了{Xt}受到的确定性影响，而反映了{Xt}受到的随机影响。

Gramer定理说明任何一个序列的波动都可以视为同时受到了确定性影响和随机性影响的综合作用。平稳时间序列要求这两方面的影响都是稳定的，而非平稳时间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方面是不稳定的。

**二、时间序列的结构形式**

非平稳时间序列(*xt*)的确定性因素分为4种：

（1）趋势变化因素(Tt)——表现出某种倾向，上升或下降或水平；

（2）季节变化因素(St)——周期固定的波动变化；

（3）循环变化因素(Ct)——周期不固定的波动变化；

（4）不规则因素(εt)——随机波动，由许多不可控的因素影响而引起的变化。

时间序列{Xt}的结构形式有三种：

（1）加法模式：*xt* = Tt+ St+ Ct+εt

（2）乘法模式：*xt* = Tt StCtεt

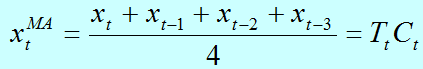
（3）混合模式：*xt* = Tt St Ct+εt

上述模式中，趋势变化Tt是基础，其它变化与趋势变化结合，构成序列{ *xt* }. 在加法模式中，各变化因素均与*xt*的单位相同；在乘法模式中，Tt与*xt*有相同的单位，其它因素的变化均数比例值；在混合模型中，Tt、εt与*xt*有相同的单位，St 和Ct是比例值。各式中的随机因素εt，均假定为独立的、方差不变的、均值为0的白噪声序列。在这些假定下，对时间序列进行分解。

**三、时间序列的传统分解法步骤**

1. 分解出长期趋势因素与循环因素

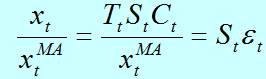
设序列的季节长度为4（一年分为4季）。由假定E(εt)=0，故只要对序列*xt*作移动长度为4的移动平均，就可消除季节和随机波动的影响（因为随机波动有正波动和负波动，一做平均，正负波动就相互抵消，随机波动影响就接近于零）。记移动平均值为：



则移动平均后的序列，即为序列的趋势因素和循环因素。类似地，若序列按月份周期，则取12。

2. 分解季节因素与随机因素

考虑乘法模式*xt*= Tt StCtεt，则两边同除以得



只含季节因素与随机因素。因此，它含有确定季节因素所必须的信息。若它的比值大于100%，就意味着序列的实际值*xt*比滑动平均值Tt Ct要大（该季度的季节性与随机性高于平均数，反之低于平均数），反之要小。

3. 从Stεt中分解季节因素St

即保留季节性，消除随机性，可以采取了按季节平均的方法，将前面得到的序列Stεt逐年逐季排列起来，然后将各年的相同季节的Stεt相加起来，再进行平均。

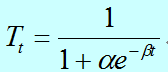
4. 从TtCt序列中分解出Ct序列

TtCt包含了趋势因素与循环因素，要把这两者分离出来，首先要确定一种能最好地描述数据的长期趋势变化的曲线类型。趋势变化曲线，可能有以下几种类型：

（1）线性趋势：Tt=a+bt

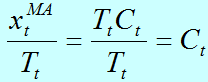
（2）指数曲线：Tt=αeβt

（3）S型曲线：



属于何种趋势曲线，要根据序列的数值进行判断，并运用最小二乘法，估计出有关参数。

确定了趋势因素Tt后，可以用下式计算出循环指数Ct：



Ct也围绕100%波动，若Ct低（高）于100%，则意味着第t年的经济活动水平低（高）于所有年份的平均水平。

**四、温特线性和季节性指数平滑**

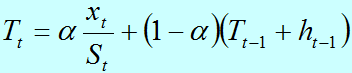
既含有线性趋势和季节性的数据进行处理和预测，使用温特（Winter）线性和季节性指数平滑方法，模型形式为：

*xt*=St(Tt+εt)

判断数据是否有季节性，粗略判断可以直接观察时序图，更好的方法是解析法，即通过研究数据序列的自相关性判断。

温特方法由三个基础的平滑公式和一个预测方程组成，每个平滑公式都含有一个平滑系数：

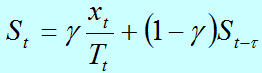
总体平滑公式：



趋势平滑公式：



季节的平滑公式：



预测公式：



其中，α, β, γ是三个不同的平滑系数，Tt是消除季节因素后的趋势平滑值，*xt*是序列的实际值，*ht*是趋势增加或减少量序列，St是季节调整因子，τ是季节的长度（如一年中的月数12或季度数4），*l*是向前预测期数，是向前*l*期的预测值。

总体平滑和趋势平滑公式是序列*xt*消除季节因素St后，霍尔特双参数α和β线性指数平滑法。季节平滑公式是序列*xt*消除趋势因素Tt后，季节指数的加权平均修匀值。以当前观察的季节指数*xt* / Tt和上期季节指数St-τ进行γ加权平均。对于乘法模型来说，季节指数围绕1波动，可能大于1，也可能小于1。在拟合模型时可以通过求解最小的均方误差MSE得到三个平滑系数的具体值。预测公式是利用拟合模型短期向前预测*l*期的预测值公式。

**（三）季节调整——PROC X11过程**

X11过程是根据美国国情调查局编制的时间序列季节调整过程X-11改编的，可以对月度或季度时间序列进行季节调整。其基本原理就是时间序列的确定性因素分解方法。

X11过程是基于这样的假定：任何时间序列都可以拆分成长期趋势波动Tt、季节波动St、不规则波动εt的影响。又有经济学家发现在经济时间序列中交易日Dt也是一个很重要的影响因素（日历天数的组成不同而引起的变动）。因此，任一时间序列可以分解乘法模型*xt*=Tt StDtεt或加法模型*xt*=Tt +St+Dt+εt。

由于宏观调控部门主要关注的是序列的长期趋势波动Tt的规律，所以X11过程主要目的是要从原序列中剔除季节影响、交易日影响和不规则波动影响，得到尽可能准确的长期趋势规律。

而采取的方法就是前文的因素剔除法和平滑技术。X11过程不依赖任何模型，普遍采用移动平均法：用多次短期中心移动平均法消除不规则波动，用周期移动平均消除趋势，用交易周期移动平均消除交易日的影响。在整个过程中总共要用到11次移动平均，所以得名为X11过程。

基本语法：

proc x11 data=数据集 </可选项> ;

monthly 选项列表;

quarterlly选项列表;

arima 选项列表;

macurves 选项;

output out=数据集 </选项列表>;

pdweights 变量

tables表名列表;

var变量列表;

by 变量;

id变量列表;

说明：

（1）monthly或quarterly语句是必不可少的，用来说明数据集是月度序列还是季度序列；

（2）pdweights和macurves语句只能与monthly语句一起用，分别用来指定星期几的权重和月份的滑动平均长度；

（3）tables语句控制各种表格的输出。output语句语句控制生成out=后指定的数据集；

（4）proc x11语句的可选项：

 outtdr＝数据集名——输出交易日回归的结果（B15表和C15表中的内容）到数据集；

 outstb＝数据集名——输出稳定季节性检验的结果（表D8中的内容）到数据集；

 outex——把在arima处理过程中预测的观察加到out=输出数据集中；

（5）arima语句及可选项

X-11方法用一系列中心化滑动平均来估计季节成分，但在起始和结尾处只能用非对称权重。非对称权重可导致季节因子估计不准，有了新数据以后就可能造成大的更改。加拿大统计局开发了一种X-11-ARIMA方法来处理该问题。使用arima语句，就是对在var语句中指定的序列应用X-11-ARIMA方法。

该方法从原始数据估计一个arima模型（使用用户指定的模型，或者通过五个预先定义的arima模型中选择一个最优的），然后用此模型把序列外推一年或几年。再根据这个延长了序列进行季节调整，此时原序列的尾部就可用对称权重了。

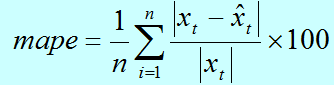
 backcast＝n——指定序列反向外推的年数，默认为0；

 chicr＝值——指定Box-Ljung拟合不足卡方检验时所用的显著水平值，默认为0.05。原假设为预定的模型（共5个）无拟合不足；

 forecast＝n——指定预报的年数，默认为1；

 mape＝值——指定平均相对误差的临界值，取值在1到100之间，默认为15。mape值作为接受还是拒绝一个模型的临界值。模型的mape值小于临界值说明模型可用，反之模型被拒绝。

mape值的计算公式如下：



其中，n=36（最后三年的月数）或12（最后三年的季度数），*xt*为原始序列的最后三年的观察值。

 maxiter＝n——指定估计过程最多允许的迭代次数，n取值为1到60之间，默认为15；

 method＝cls | uls | ml——指定估计方法，分别为条件最小二乘法、无条件最小二乘法、最大似然估计；

 model＝(P=n1 Q=n2 SP=n3 SQ=n4 DIF=n5 SDIF=n6)——指定arima模型。P和SP表示一般的和季节的自回归过程（AR）阶数；Q和SQ表示一般的和季节的移动平均过程（MA）阶数，DIF和SDIF表示一般的和季节的差分阶数；季节s=12（对应monthly）或4（对应quarterly）。

例如，指定一个(0,1,1)(0,1,1)s模型，表示(P,DIF,Q)(SP,SDIF,SQ)s模型。假设考虑月度序列s=12，且E(xt)=μ，则具体模型形式为：





 ovdifcr＝值——指定对5个预先定义模型进行过度差分检验时所用的临界值。取值范围在0.8到0.99之间，默认为0.9。五个模型都有一个季节MA因子，最多两个非季节因子（模型2、4、5）。有季节差分和非季节差分。以模型2例，那么具体模型形式为：



若θ3=1，则等式两边可以消去(1-B12)项，得到低阶模型。类似地，如果θ1+θ2=1，则又可以消去(1-B)项，得到低阶模型。因为参数估计必定有误差，要求小于1是不合理的。因此，过度差分检验的要求为：

θ3≤0.9 且 θ1+θ2≤0.9

大于0.9应拒绝此模型。

 transform＝(log) | (a\*\*b)——允许在对模型进行估计之前先进行用户指定的一些变换，产生预报值后再变换回原来的取值。(log)是自然对数变换，(a\*\*b)是乘方变换：*xt*=(*xt*+*a*)*b*.

（6）macurves语句

该语句只适用于月度数据，为任一月份指定估计季节因子：月份=选项值。例如：

macurves jan=’3’ feb=’3x5’ mar=stable;

’3’——3期移动平均；

’3X3’、 ’3X5’、’3X9’——3×3、3×5（移动平均,5期移动平均再做3次移动平均）3×9移动平均；

’stable’——所有值的平均值作为恒定的季节因子；

（7）monthly语句

月度时间序列数据集必须使用monthly语句。主要选项为：

 additive——指定进行加法模型季节调整。默认为乘法模型；

 charts＝standard | full | none——指定生成的图表类型。默认为standard，生成12月度季节性图表和趋势起伏图表；full选项，还额外输出不规则项和季节因子的图表；none选项，不输出任何图表；

data＝日期变量名 start=mmmyy end＝mmmyy——指定要处理的部分时间序列数据的起止时间，例如：

monthly date=date start=jan90 end=dec99;

exclude=值——在交易日回归时把偏离均值超过指定值倍数的标准差的不规则值排除在外。取值在0.1到9.9之间，默认为2.5；

pmfactor=月份因素变量——用于调整已经知道特殊原因的月份数据，例如，某公司2005年1月份罢工，销售额sales比往常下降了约50%，这是一个原因已知的一次性事件，应该预先修正该月份的销售额，才能排除罢工的影响。在原时间序列数据集中设置一个反映月份因素的新变量x，其他月份的新变量x值都设定为100，即sales=sales/x×100=sales（销售额不用调整）；2005年1月份的新变量值x设定为50，即sales=sales/x×100=2×sales，销售额还原成经验值，示例：

monthly date=date pmfactor=x;

 fullweight=值——设定观察值距离均值小于指定值倍数的标准差，将赋予观察值的权数为最大值1。缺省值为1.5。

 zeroweight=值——设定观察值距离均值大于指定选项值倍数的标准差，将赋予观察值的权数为最小值0。缺省值为2.5。选项zeroweight=的值必须大于选项fullweight=的值。观察值距离均值落入fullweight=值和zeroweight=值之间，将被赋予0到1之间的一个线性分级的权重值。

 printout=standard | long | full | none——指定打印哪些表格。

（8）quarterly语句

季度时间序列数据集必须使用quarterly语句，其主要选项和用法与monthly类似。季度时间值的格式为：1999年第一季度为’99Q1’。

（9）pdweights语句

用来指定星期一到星期七的权重值，只能用于月度数据。选项格式为：星期几=权重值。这些权重值是用来计算先验交易日因子，而先验交易日因子是在季节调整过程之前对原始序列进行修正的。只需给出相对权重，X11过程会自动调整到相加之和为7。例如：

pdweights sun=0.1 mon=0.9 tue=1 wed=1 thu=1 fri=0.7 sat=0.3;

（10）tables语句

tables语句用来指定打印一些额外表格。例如，如果省略选项printout=，下面语句只打印最终季节因子和最终季节调整过的序列。

tables d10 d11;

（11）output语句

用来生成包含指定表格的输出数据集，输出数据集名由选项out=给出。对每一张要进入输出数据集的表格，由选项：表格名=新变量名列表，来指定。下面是一个var语句和output语句的示例：

var z1 z2 z3;

output out=out\_x11 b1=x1 d11=t1 t2 t3;

首先var语句指定输入数据集中三个数值型变量z1、z2和z3分别进行季节调整过程分析。选项b1=x1指定对变量z1进行分析，结果b1表格存入到新变量x1中；选项d11=t1 t2 t3指定对三个数值变量z1、z2、z3进行分析，三个结果b11表格分别存入到新变量t1、t2、t3中。

**例1**对1993－2005年中国社会消费品96个月份零售总额的时间序列数据：

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1993年 | 1994年 | 1995年 | 1996年 | 1997年 | 1998年 | 1999年 | 2000年 |
| 1月 | 977.5 | 1192.2 | 1602.2 | 1909.1 | 2288.5 | 2549.5 | 2662.1 | 2774.7 |
| 2月 | 892.5 | 1162.7 | 1491.5 | 1911.2 | 2213.5 | 2306.4 | 2538.4 | 2805.0 |
| 3月 | 942.3 | 1167.5 | 1533.3 | 1860.1 | 2130.9 | 2279.7 | 2403.1 | 2627.0 |
| 4月 | 941.3 | 1170.4 | 1548.7 | 1854.8 | 2100.5 | 2252.7 | 2356.8 | 2572.0 |
| 5月 | 962.2 | 1213.7 | 1585.4 | 1898.3 | 2108.2 | 2265.2 | 2364.0 | 2637.0 |
| 6月 | 1005.7 | 1281.1 | 1639.7 | 1966.0 | 2164.7 | 2326.0 | 2428.8 | 2645.0 |
| 7月 | 963.8 | 1251.5 | 1623.6 | 1888.7 | 2102.5 | 2286.1 | 2380.3 | 2597.0 |
| 8月 | 959.8 | 1286.0 | 1637.1 | 1916.4 | 2104.4 | 2314.6 | 2410.9 | 2636.0 |
| 9月 | 1023.3 | 1396.2 | 1756.0 | 2083.5 | 2239.6 | 2443.1 | 2604.3 | 2854.0 |
| 10月 | 1051.1 | 1444.1 | 1818.0 | 2148.3 | 2348.0 | 2536.0 | 2743.9 | 3029.0 |
| 11月 | 1102.0 | 1553.8 | 1935.2 | 2290.1 | 2454.9 | 2652.2 | 2781.5 | 3108.0 |
| 12月 | 1415.5 | 1932.2 | 2389.5 | 2848.6 | 2881.7 | 3131.4 | 3405.7 | 3680.0 |

使用X11过程进行季节调整，假设先不考虑日历效应和不需要对数据进行任何预先的调整。由于没有交易日的影响，我们考虑使用乘法模型*xt*= Tt Stεt.

代码：

**data** sales;

input sales @@;

date = intnx( 'month', **'01jan1993'd**, \_n\_-**1** );

format date monyy5.;

datalines;

977.5 892.5 942.3 941.3 962.2 1005.7 963.8 959.8 1023.3 1051.1 1102 1415.5

1192.2 1162.7 1167.5 1170.4 1213.7 1281.1 1251.5 1286 1396.2 1444.1 1553.8 1932.2

1602.2 1491.5 1533.3 1548.7 1585.4 1639.7 1623.6 1637.1 1756 1818 1935.2 2389.5

1909.1 1911.2 1860.1 1854.8 1898.3 1966 1888.7 1916.4 2083.5 2148.3 2290.1 2848.6

2288.5 2213.5 2130.9 2100.5 2108.2 2164.7 2102.5 2104.4 2239.6 2348 2454.9 2881.7

2549.5 2306.4 2279.7 2252.7 2265.2 2326 2286.1 2314.6 2443.1 2536 2652.2 3131.4

2662.1 2538.4 2403.1 2356.8 2364 2428.8 2380.3 2410.9 2604.3 2743.9 2781.5 3405.7

2774.7 2805 2627 2572 2637 2645 2597 2636 2854 3029 3108 3680

;

**run**;

**proc** **x11** data=sales;

monthly date=date;

var sales;

arima maxit=**60**;

tables d11;

output out=out b1=series d10=season d11=adjusted d12=trend d13=irr;

**proc** **print** data=out;

**run** ;

title 'Monthly Retail Sales Data';

**proc** **sgplot** data=out;

series x=date y=series / markers

markerattrs=(color=red symbol='asterisk')

lineattrs=(color=red)

legendlabel="original" ;

series x=date y=adjusted / markers

markerattrs=(color=blue symbol='circle')

lineattrs=(color=blue)

legendlabel="adjusted" ;

yaxis label='Original and Seasonally Adjusted Time Series';

**run**;

title 'Monthly Seasonal Factors (in percent)';

**proc** **sgplot** data=out;

series x=date y=season / markers markerattrs=(symbol=CircleFilled) ;

**run**;

title 'Monthly Retail Sales Data (in $1000)';

**proc** **sgplot** data=out;

series x=date y=trend / markers markerattrs=(symbol=CircleFilled) ;

**run**;

title 'Monthly Irregular Factors (in percent)';

**proc** **sgplot** data=out;

series x=date y=irr / markers markerattrs=(symbol=CircleFilled) ;

**run**;

运行结果及说明：

日期变量date从intnx()函数获得从1993年1月1日开始每过一个月的时间。

intnx()函数有3个参数：参数1是指定等时间间隔’month’，还可以取’day’、’week’、’quarter’、’year’等；参数2指定参照时间’01jan1993’；参数3是指定开始的时间指针\_n\_k，k为整数。k取正值（负值），开始时间为参照时间向未来（过去）拨k期。

调用季节调整X11过程之前，应该先绘制原始时间序列的散点图（略，见后面原始序列与调整序列对比图），直观判断一下是否存在确定性季节波动，以便确定能否调用X11过程。如果的确存在季节性波动，还需要判断一下季节性的时间周期为月份还是季节。

本例是月度数据，必须要用monthly date=date语句。

X11 过程

季节调整 - sales

|  |
| --- |
| X-11 季节调整程序 |
| 美国人口普查局 |
| 经济研究和分析部门 |
| 1968 年 11 月 1 日 |
|  |
| X-11 程序分为七大部分。 |
| 部分 说明 |
| A. 先验调整（若有） |
| B. 不规则成分权重 |
| 和回归交易日因子的初步估计值 |
| C. 上述统计量的最终估计值 |
| D. 季节、趋势周期和不规则成分 |
| 的最终估计值 |
| E. 分析表 |
| F. 汇总测度 |
| G. 图 |
|  |
| 序列 - sales |
| 涉及的期间 - 1/1993 至 12/2000 |
| 运行的类型: 乘法季节调整。 |
| Selected Tables or Charts. |
| 毕业极值的 Sigma 限制是 1.5 和 2.5 |
| 已排除那些超出 2.5-sigma 限制的不规则值 |
| 从交易日回归 |

概述表

arima语句的作用是把时间序列延长，使得序列尾部可以使用对称移动平均方法，用以解决减少对序列尾部的更正。对时间序列延长的模型，从五个预先定义的模型中择优采用（也可以用 model=来自己定义）。参数maxit=60指定估计过程最多允许迭代60次，特别是对于高阶arima模型，缺省值最多允许迭代15次可能不够。

| **Conditional Least Squares Estimation** | | | | |
| --- | --- | --- | --- | --- |
| **参数** | **估计值** | **近似标准误差** | **t 值** | **滞后** |
| **MU** | 0.6258994 | 4.4939186 | 0.14 | 0 |
| **MA1,1** | 0.0511338 | 0.2597677 | 0.20 | 1 |
| **MA1,2** | -0.551101 | 0.1492533 | -3.69 | 2 |
| **MA2,1** | 0.0570263 | 0.1231064 | 0.46 | 12 |
| **AR1,1** | -0.515567 | 0.2732939 | -1.89 | 1 |
| **AR1,2** | -0.444158 | 0.2026064 | -2.19 | 2 |

| **Conditional Least Squares Estimation** | | |
| --- | --- | --- |
| **Variance Estimate =** | 3164.8385 |  |
| **Std Error Estimate =** | 56.256897 |  |
| **AIC =** | 910.28403 | \* |
| **SBC =** | 924.79708 | \* |
| **Number of Residuals=** | 83 |  |
| **\* Does not include log determinant** | | |

|  |
| --- |
| 模型 5 的条件汇总: (2,1,2)(0,1,1)s, 无转换 |
|  |
| Box-Ljung 卡方: 17.51，自由度: 19，概率: 0.56 |
| （条件概率 > 0.05） |
| 过差分检验: MA 参数的总和 = 0.06 |
| （必须小于 0.90） |
| MAPE - 前三年: 1.49 （必须小于 15.00 %） |
| - 去年: 1.37 |
| - 前年: 1.76 |
| - 大前年: 1.34 |

调用arima语句时自动从五个预先定义的模型中选择的最优模型，输出结果表明选择了模型5：(2,1,2)(0,1,1)s，并给出了模型参数的估计值。可得arima模型的具体表达式为：



用它延长1年（12个月）的时间序列值。对此模型残差进行拟合不足的Box-Ljung 卡方检验，卡方值为17.51 ，自由度为 19 ，相应概率为 0.56>0.05，不能拒绝模型残差拟合充分的原假设。

对此模型过度差分检验，二阶MA参数值之和为0.06（0.0570263）小于0.9标准，因此，该模型不存在过度差分问题。对该模型用最后三年的原始序列来检验平均相对误差MAPE准则，计算结果为1.49 %小于临界值15.00 %，因此该模型的误差是在可以接受的范围内。

最后要特别注意，此模型许多参数值的t检验并没有通过，t值太小，并不能拒绝这些参数值为0的原假设，例如MU、MA1,1和MA2,1参数值t检验所计算的t值。另外从直观上也可以看出这些参数值都很小（接近0）。但该模型已经是五个预先定义模型中最优的。

tables d11语句指定打印d11表格，它输出最终的季节调节后的序列。output语句把部分结果输出到out数据集，表b1中的原序列值输出到series，表d10中的最终季节因子输出到season，表d11中的最终季节调节后的序列值输出到adjusted，表d12中的最终趋势起伏值输出到trend，表d13中的最终不规则序列值输出到irr.

| **D11 最终季节调整序列** | | | | | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **年** | **1 月** | **2 月** | **3 月** | **4 月** | **5 月** | **6 月** | **7 月** | **8 月** | **9 月** | **10 月** | **11 月** | **12 月** | **合计** |
| **1993** | 938.400 | 903.801 | 978.115 | 995.791 | 1012.94 | 1035.25 | 1038.59 | 1040.35 | 1046.92 | 1059.10 | 1054.33 | 1111.46 | 12215.1 |
| **1994** | 1143.79 | 1175.13 | 1212.12 | 1238.63 | 1278.79 | 1320.14 | 1349.55 | 1394.30 | 1428.09 | 1451.94 | 1484.47 | 1517.64 | 15994.6 |
| **1995** | 1536.45 | 1503.37 | 1594.11 | 1641.30 | 1674.23 | 1695.29 | 1752.78 | 1773.78 | 1794.01 | 1818.97 | 1842.13 | 1877.89 | 20504.3 |
| **1996** | 1830.04 | 1921.80 | 1939.16 | 1970.40 | 2012.47 | 2042.90 | 2040.78 | 2075.36 | 2126.95 | 2135.94 | 2173.96 | 2245.15 | 24514.9 |
| **1997** | 2189.00 | 2217.93 | 2225.13 | 2237.45 | 2241.45 | 2261.39 | 2275.99 | 2277.50 | 2281.56 | 2316.37 | 2323.00 | 2279.17 | 27125.9 |
| **1998** | 2435.51 | 2306.03 | 2383.43 | 2406.08 | 2413.42 | 2442.45 | 2477.85 | 2504.36 | 2485.02 | 2482.45 | 2505.42 | 2487.64 | 29329.7 |
| **1999** | 2539.83 | 2534.06 | 2512.44 | 2522.47 | 2519.49 | 2561.00 | 2583.12 | 2606.18 | 2643.11 | 2667.78 | 2621.99 | 2713.58 | 31025.0 |
| **2000** | 2648.64 | 2800.86 | 2748.74 | 2757.17 | 2810.39 | 2797.94 | 2819.96 | 2848.87 | 2893.17 | 2932.54 | 2928.56 | 2939.14 | 33926.0 |
| **Avg** | 1907.71 | 1920.37 | 1949.15 | 1971.16 | 1995.40 | 2019.54 | 2042.33 | 2065.09 | 2087.36 | 2108.13 | 2116.73 | 2146.46 |  |

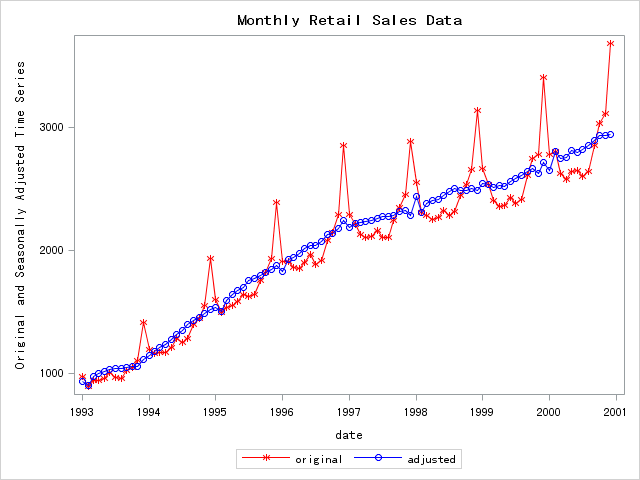
合计: 194635 均值: 2027.5 S.D.: 595.95

最终季节调整后每月的销售额、每月销售额的总和、每月销售额的平均值和标准差。

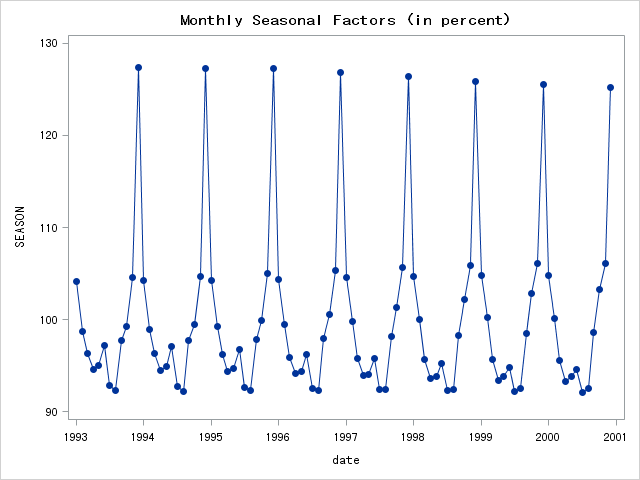
|  |
| --- |
| Monthly Retail Sales Data（部分） |

| **Obs** | **date** | **SERIES** | **SEASON** | **ADJUSTED** | **TREND** | **IRR** |
| --- | --- | --- | --- | --- | --- | --- |
| **1** | JAN93 | 977.5 | 104.167 | 938.40 | 935.79 | 100.279 |
| **2** | FEB93 | 892.5 | 98.750 | 903.80 | 954.38 | 94.700 |
| **3** | MAR93 | 942.3 | 96.338 | 978.11 | 974.71 | 100.350 |
| **4** | APR93 | 941.3 | 94.528 | 995.79 | 995.74 | 100.005 |
| **5** | MAY93 | 962.2 | 94.991 | 1012.94 | 1015.29 | 99.769 |
| **6** | JUN93 | 1005.7 | 97.146 | 1035.25 | 1029.16 | 100.591 |

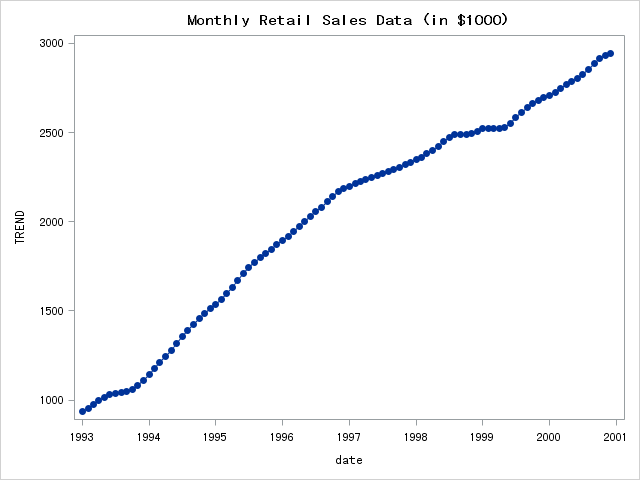
输出数据集out.



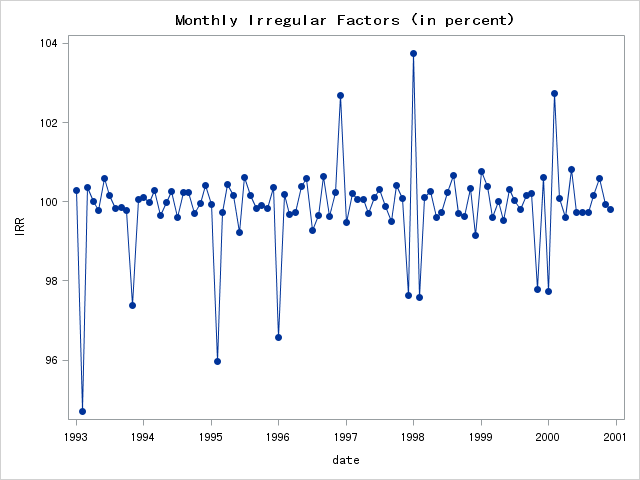
原序列与消除季节效应后的调整序列对比图



使用多次移动平均和迭代方法求出的最终季节因子，最终季节因子好象在缓慢地减小，而在原始序列中却没有这么明显。



零售额剔除季节效应之后，使用移动平均方法拟合的序列趋势，有非常显著的线性递增趋势。



消除了季节和趋势因素后，残差序列很不规则，说明季节和趋势信息提取很充分，因此，用X11过程来拟合中国社会消费品零售总额序列的把握还是比较准确的。