时间序列分析 2

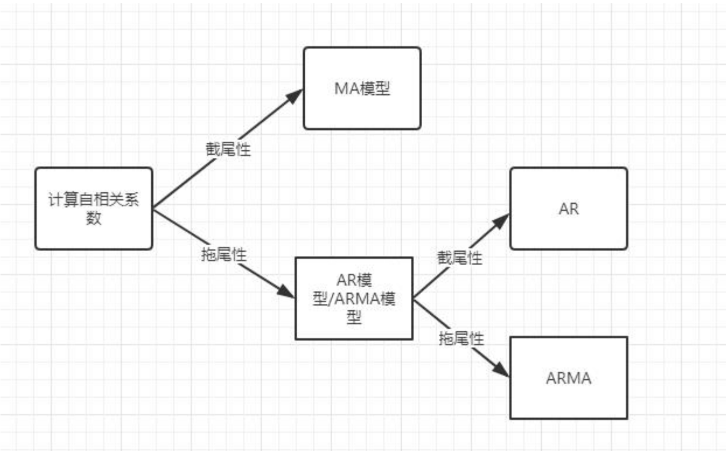
继去趋势后的自/偏相关值估计

2016100104028　李科

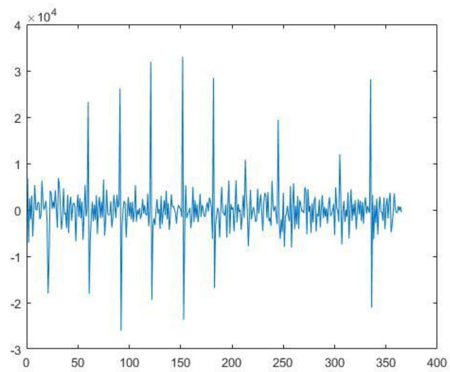
内容简述：利用去趋势和去周期后的移动开户数据，对数据进行进一步的分析。

1. 再次验证处理后数据已经是平稳时间序列，且由自相关系数的截尾性初步判断可用 MA(1) 模型。
2. 去均值使该时间序列变为零均值平稳过程。
3. 自己编写函数实现计算**无偏/有偏**自相关系数的功能并与 matlab 自带 autocorr 进行比较。
4. 自己编写函数根据自相关系数，这里采用迭代而非直接求解计算偏相关函数。用计算出的部分值与 matlab 自带 parcorr 进行比较。
5. 根据自相关系数和偏相关系数的收敛趋势，得出结论：
   5.1 有偏自相关系数估计收敛速度较无偏快
   5.2 偏相关函数具有拖尾性。

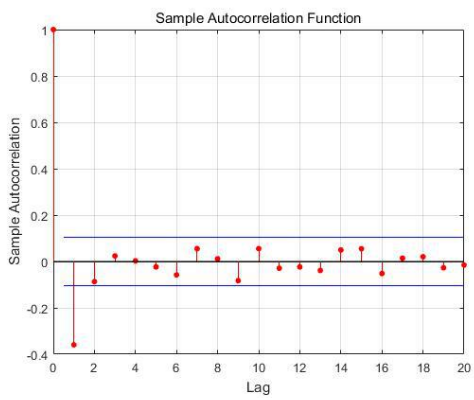
适用模型判别：



1.去趋势和去周期后数据的平稳性



图一　处理后的数据



图二　处理后数据的 20 期自相关系数

自相关系数定义为:

$$\rho(h)=\frac{E[X_tX_{t+h}]}{\sqrt{EX_t^2}\sqrt{EX_{t+h}^2}}=\begin{cases}\sum_{j=0}^{q}\theta_j\theta_{j+|h|}/\sum_{j=0}^{q}\theta_j^2, & |h|\le q;\\ 0, & |h|>q\end{cases}$$

由图二观察知道，当|h|>1 时，自相关系数为近似为 0.认为自协方差函数(自相关函数)有“截尾性”.可以使用 MA(1)模型。为了考虑能否使用 ARMA 模型，进一步进行讨论。

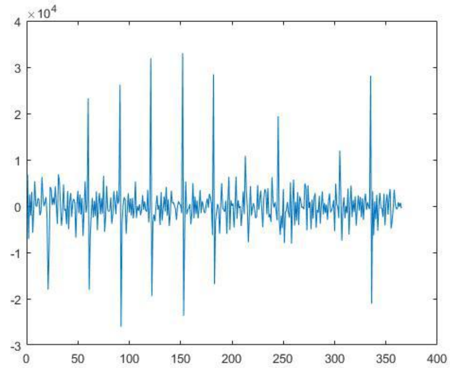
2. 转换为零均值平稳序列

采用 mean 函数计算该序列的均值，处理如下:
随机变量总体 X 的样本的均值估计量:

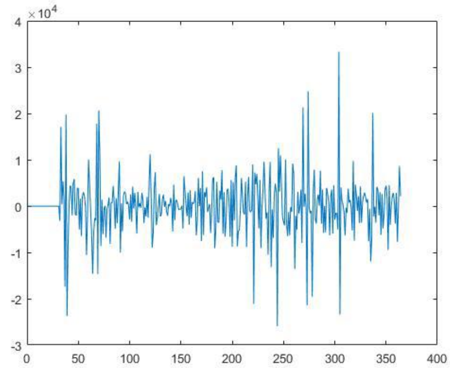
$$\overline{X}_n=\frac{1}{n}\sum_{k=1}^{n}X_k$$

```
%% 平稳序列去均值变为零均值平稳序列
zeromean_detrend_deT_data = detrend_deT_data-mean(detrend_deT_data);
```

转换前后序列变化如下:



图三 非零均值平稳序列



图四 零均值平稳序列

3. 无偏/有偏自相关函数的估计

自协方差函数（或自相关函数）的样本估计量通常有两种类型:

$$\begin{cases}\hat{\gamma}(h)=\hat{\gamma}(-h)=\frac{1}{N}\sum_{t=1}^{N-h}X_tX_{t+h}\\ \hat{\rho}(h)=\hat{\gamma}(h)/\hat{\gamma}(0), \qquad h=0,1,2,\cdots\end{cases} \quad (3.3.1)$$

$$\begin{cases}\hat{\gamma}(h)=\hat{\gamma}(-h)=\frac{1}{N-h}\sum_{t=1}^{N-h}X_tX_{t+h}\\ \hat{\rho}(h)=\hat{\gamma}(h)/\hat{\gamma}(0), \qquad h=0,1,2,\cdots\end{cases} \quad (3.3.2)$$

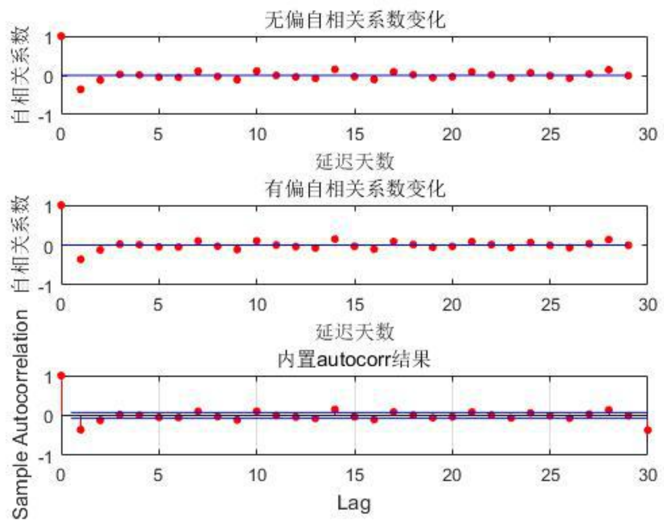
关于两种估计量,前人曾经做过比较，G.M.Jenkins 结论：
1.估计量（ 3.3.1） 更满意。
2.其中（3.3.1）是自相关函数的有偏估计，（3.3.2）是自相关函数的无偏估计。
3.（3.3.1）式定义的估计量收敛于零的速度更快。
4.由(3.3.1)式定义的样本自协方差函数矩阵（或相关矩阵）是正定的

根据两种自相关系数的计算公式，编写函数如下：
其中函数输入为零均值的平稳序列，输出为一组无偏估计的自相关系数，一组为有偏的自相关系数。完成自己的函数编写后，选择计算前面 30 期自相关系数与 matlab 自带 autocorr 进行比较。

自定义计算无偏和有偏两种自相关系数代码：
```
function [unbiased_autocorr,biased_autocorr] =
my_autocorr(time_vector)
%% 函数定义,计算无偏和有偏两种自相关系数
n = length(time_vector);
unbiased_autocorr = zeros(n,1);
biased_autocorr = zeros(n,1);
% 延迟期数从 0 到 n
for h = 0:n
    omega = 0;
    % 累加从 t=1 到 n-h
    for t = 1:n-h
        omega = omega+time_vector(t)*time_vector(t+h);
    end
    unbiased_autocorr(h+1) = omega/(n-h);
    biased_autocorr(h+1) = omega/n;
end
% 自相关函数转换为自相关系数
unbiased_autocorr = unbiased_autocorr/unbiased_autocorr(1);
biased_autocorr = biased_autocorr/biased_autocorr(1);
%% 绘图比较
n_compare = 30;
figure()
subplot(311)
plot(0:n_compare-1,unbiased_autocorr(1:n_compare),'ro','MarkerFaceCol
or','r','markersize',4)
hold on,plot(0:n_compare-1,zeros(1,n_compare),'b')
title('无偏自相关系数变化'),xlabel('延迟天数'),ylabel('自相关系数')
subplot(312)
plot(0:n_compare-1,biased_autocorr(1:n_compare),'ro','MarkerFaceColor
','r','markersize',4)
hold on,plot(0:n_compare-1,zeros(1,n_compare),'b')
```

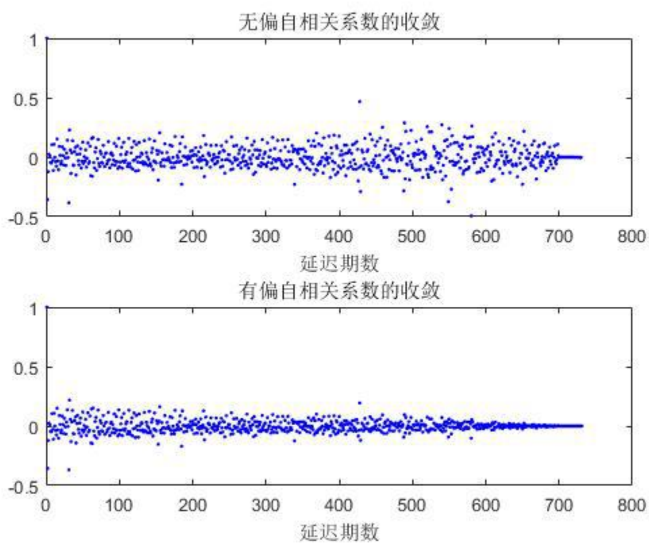
```
title('有偏自相关系数变化'),xlabel('延迟天数'),ylabel('自相关系数')
subplot(313)
autocorr(time_vector,n_compare),title('内置 autocorr 结果')
```

结果如图：



图五　前 30 期自相关系数比较图

由图，计算结果与内置函数结果十分接近，验证了自己构建的函数的准确性。



图六 无偏/有偏自相关系数的收敛速度比较

可以看出。该图验证了第三个结论，即有偏估计**定义的估计量收敛于零的速度更快，且从收敛趋势看来效果更好。**

4. 偏相关系数的估计和比较

计算偏相关系数有两种方法，直接求解和递推求解。这里考虑递推求解。运用 3 求解得到的自相关系数一步步递推得到所有的偏相关系数。

4.1 直接求解

$$\phi_{kk}=\frac{\gamma_k-\alpha_1\gamma_{k-1}-\alpha_2\gamma_{k-2}-\cdots-\alpha_{k-1}\gamma_1}{\gamma_0-\alpha_1\gamma_1-\alpha_2\gamma_2-\cdots-\alpha_{k-1}\gamma_{k-1}}=\frac{\rho_k-\alpha_1\rho_{k-1}-\alpha_2\rho_{k-2}-\cdots-\alpha_{k-1}\rho_1}{\rho_0-\alpha_1\rho_1-\alpha_2\rho_2-\cdots-\alpha_{k-1}\rho_{k-1}} \quad (3.4.1)$$

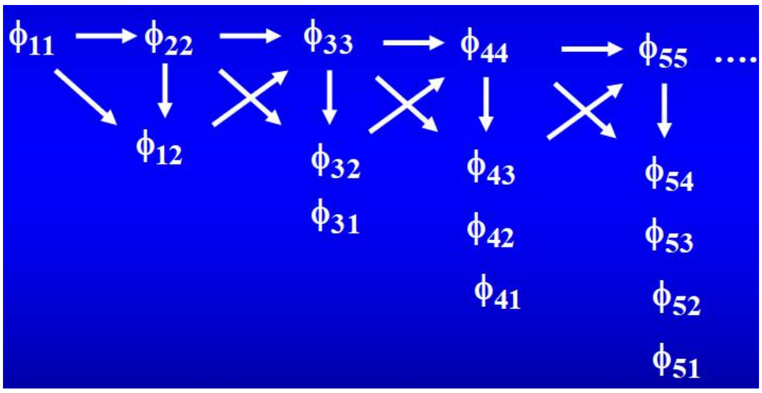
由最优线性估计的计算公式，应用矩阵按行（列）展开式的性质可得

$$\phi_{kk}=\frac{\begin{vmatrix}1&\rho_1&\rho_2&\cdots&\rho_{k-2}&\rho_{k-1}\\ \rho_1&1&\rho_1&\cdots&\rho_{k-3}&\rho_{k-2}\\ \vdots&\vdots&\vdots&\ddots&\vdots&\vdots\\ \rho_{k-1}&\rho_{k-2}&\rho_{k-3}&\cdots&\rho_1&\rho_k\end{vmatrix}}{\begin{vmatrix}1&\rho_1&\rho_2&\cdots&\rho_{k-2}&\rho_{k-1}\\ \rho_1&1&\rho_1&\cdots&\rho_{k-3}&\rho_{k-2}\\ \vdots&\vdots&\vdots&\ddots&\vdots&\vdots\\ \rho_{k-1}&\rho_{k-2}&\rho_{k-3}&\cdots&\rho_1&1\end{vmatrix}} \quad (3.4.2)$$

4.2 递推求解

$$\begin{cases}\phi_{11}=\rho(1),\\ \phi_{k+1,k+1}=[\rho(k+1)-\sum_{j=1}^{k}\rho(k+1-j)\phi_{kj}][1-\sum_{j=1}^{k}\rho(j)\phi_{kj}]^{-1},\\ \phi_{k+1,j}=\phi_{kj}-\phi_{k+1,k+1}\phi_{k,k+1-j}, \qquad (j=1,2,\cdots,k)\end{cases} \quad (3.4.3)$$

递推过程：

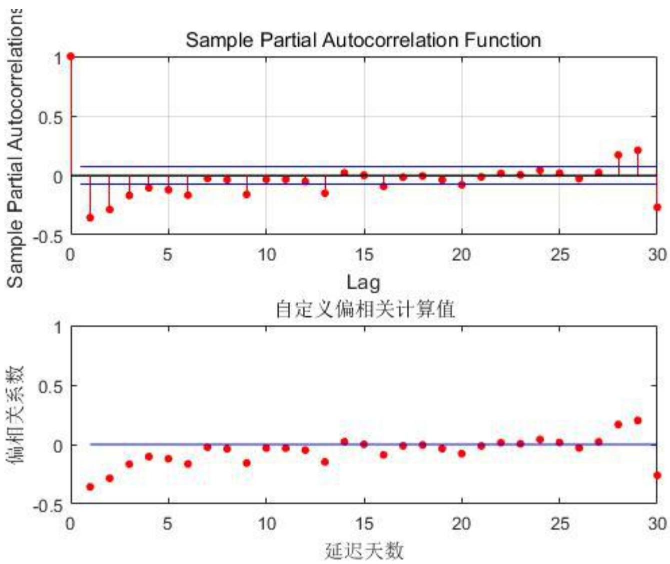


图七 递推求解过程

理解以上递推思想后，利用 3 计算出来的自相关系数，编写函数如下:
其中输入为自相关系数和之前的零均值序列。前者用于递推，后者用于利用 matlab 自带函数 parcorr 对自定义的函数计算值进行检验。
代码：

```
function [customed_parcorr] = my_parcorr(time_vector,auto_corr)
%% 函数定义
n = length(auto_corr);
customed_parcorr = zeros(n,n);
customed_parcorr(1,1) = auto_corr(2);
for k = 1:n-2
    sum1 = 0;
    sum2 = 0;
    for j = 1:k
        sum1 = sum1 + auto_corr(k+2-j)*customed_parcorr(k,j);
        sum2 = sum2 + auto_corr(j+1)*customed_parcorr(k,j);
    end
    customed_parcorr(k+1,k+1) = (auto_corr(k+2)-sum1)/(1-sum2);
    for j = 1:k
        customed_parcorr(k+1,j) = customed_parcorr(k,j) -
customed_parcorr(k+1,k+1)*customed_parcorr(k,k+1-j);
    end
end
%% 绘图比较
a = diag(customed_parcorr);
subplot(211)
compare_n = 40;
parcorr(time_vector,compare_n)
subplot(212)
plot(1:compare_n,a(1:compare_n),'ro','markerfacecolor','r','markersiz
e',4)
hold on,plot(1:compare_n,zeros(1,compare_n),'b')
ylim([-0.5,1])
title('自定义偏相关计算值')
xlabel('延迟天数'),ylabel('偏相关系数')
```

结果如图八所示:



图八 自定义偏相关系数计算值与 parcorr 前 30 期计算值比较

由比较图看来，计算结果相当。

由图八可知，偏相关系数具有拖尾性。

附录

1 主函数 detrend_main.m

```
clearvars
%% 读取数据并预处理
open_nums = xlsread('移动通知户开户数.xlsx',1,'B2:B732');
%% 计算月平均，季度平均，年平均并绘图找周期规律,并处理异常值
figure(1),[months_mean,seasons_mean,years_mean,month_starts,month_end
s,season_starts,season_ends] = Calu_mean(open_nums);
[m,n] = find(abs(open_nums-mean(open_nums))>2*std(open_nums));
open_nums(m) = mean(open_nums);
%% 尝试进行去趋势和去周期,考虑到存在多重周期，将年份分开按照月为周期去除周期（其中
每月为多少天根据两年分别不同）
data_2012 = open_nums(1:366);
data_2013 = open_nums(367:end);
figure(2),[detrend_data2012,detrend_deT_data2012] =
Detrend_plot(data_2012,30);% 2012 开户数变化趋势
figure(3),[detrend_data2013,detrend_deT_data2013] =
Detrend_plot(data_2013,31);% 2013 开户数变化趋势
figure(4),[detrend_data,detrend_deT_data] =
Detrend_plot(open_nums,30);% 2012-2013 开户数变化趋势
figure(5),subplot(411),autocorr(open_nums),title('原数据自相关图')% 自相
关图分析
subplot(412),autocorr(detrend_data),title('2012 年随机项自相关图')
```

```
subplot(413),autocorr(detrend_deT_data2013),title('2013 年随机项自相关图
')
subplot(414),autocorr(detrend_deT_data),title('2012-2013 年随机项自相关
图')
clearvars month_starts month_ends season_starts season_ends data_2012
data2013
%% 平稳序列去均值变为零均值平稳序列
zeromean_detrend_deT_data = detrend_deT_data-mean(detrend_deT_data);
[unbiased_autocorr,biased_autocorr] =
my_autocorr(zeromean_detrend_deT_data);
%% 计算偏相关函数
[customed_parcorr] =
my_parcorr(zeromean_detrend_deT_data,biased_autocorr);
```