

笔试答卷

侯怡枫

2017 年 2 月 9 日

答卷中用到的源代码都将在 Email 附件中附上。所有程序都在 Linux Ubuntu LTS 16.04 操作系统上编译。

第 1 题

这一题的源代码在 ans1.R 文件中。首先用 R 将数据 plot 出来获得一个直观的感受。

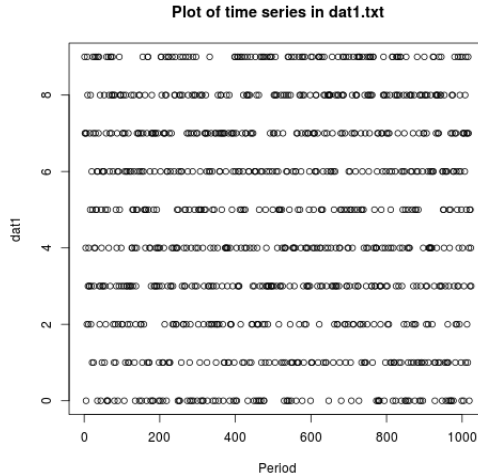


图 1: dat1.txt 中数据的时间序列图

根据图 1，改时间序列并不呈现确定性时间趋势（deterministic time trend）。然后，图 2 展示改题数据的样本 ACF。图 3展示改题数据的样本 PACF。

从图 2中可以看到，该序列和 $k = 1$ 的滞后项的样本相关系数接近 1。而大于 1 的滞后项的相关系数则呈指数化递减。从图 3中可以看到，偏自相关函数在 0 附近浮动。这显示此序列可能由一个自回归（AR）过程产生。

答题者选择“存在单位根”作为原假设。相应的备择假设是“产生改序列的随机过程是平稳的”。答题者计划使用 ADF（augmented Dickey-Fuller）检验。这里假设该序列是由一个自回归（AR）模型产生。为使用 ADF 检验，首先要选择一个信息标准（Information Criterion）来确定自

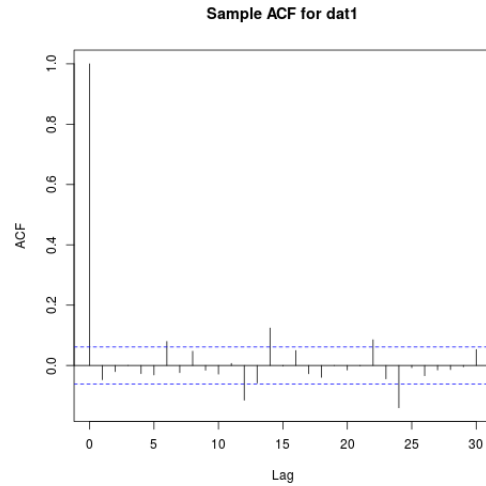


图 2: dat1.txt 中数据的样本 ACF plot

回归的滞后步长 (lag)。这里答题者选择 AIC 进行模型选择。对本题时间序列运行 AIC 后确定的滞后步长是 24。

在确定滞后步长后，运行 ADF 检验后得到的 Dickey-Fuller 统计值是-7.6675，对应的 p 值小于 0.01。所以结论是拒绝“存在单位根的原假设”。

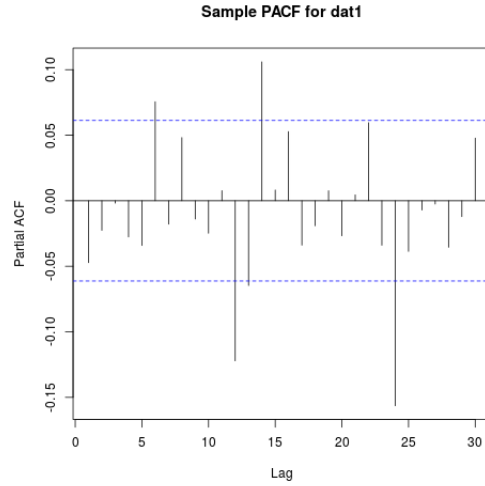


图 3: dat1.txt 中数据的样本 PACF plot

第四题

Hurst 是用来度量时间序列的长期记忆性。Hurst 指数的值在 0.5 到 1 之间意味着长期的正相关性，简单来说意味着序列中一个较高的值往往会跟随者另一个较高的值。Hurst 指数的值在 0 到 0.5 之间意味着长期的负相关性。Hurst 指数的值等于 0.5 则意味着一个完全不相关的序列。Hurst 指数的定义如下：

$$E\left[\frac{R(n)}{S(b)}\right] = Cn^H, n \rightarrow \infty \quad (1)$$

其中 $R(n)$ 是前 n 个值的值域， $S(n)$ 是标准差， $E[x]$ 是期望值， n 是观察的期数， C 是一个常数。