X13-seats 季节性调整与春节节日效应

贾昊天

中国国际金融股份有限公司

Haotian.Jia@cicc.com.cn

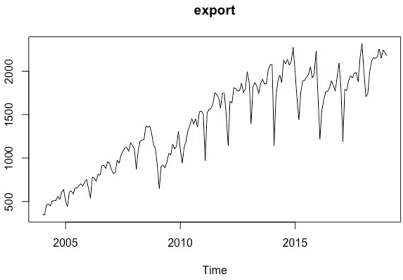
March 8, 2019

Overview

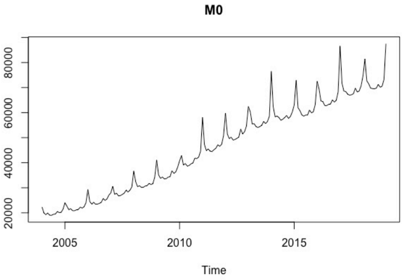
1. 研究季调的意义
2. 季节性调整的原理
3. 春节效应原理
   - 流量数据的春节效应
   - 存量数据的春节效应
4. 模型选择与参数设定
   - 模型选择
   - 参数设定

研究季调的意义

- 许多宏观经济数据存在显著的季节性与周期性，在春节时期效应尤其明显



(a) 出口



(b) M0

季节性调整的原理

- x-13-seats 使用的 regARIMA 模型，其具体的表达式如下：

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D(y_t - \sum_i \beta_i x_{it}) = \theta(B)\Theta(B^s)a_t$$

做一个简单的变形，该模型等价于

$$(1-B)^d(1-B^s)^D y_t = \sum_i \beta_i (1-B)^d(1-B^s)^D x_{it} + w_t$$

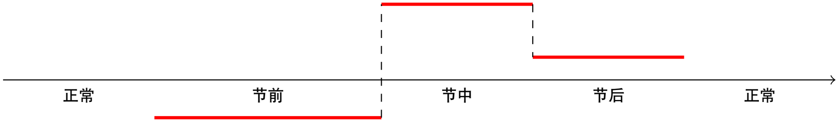
春节效应原理

- 春节效应分为节前节中和节后，数据分为流量和存量
  - 节前，节中节后表示春节效应的影响范围，即节前多少天，节中多少天，节后多少天的经济数据受到春节节日效应的影响
  - 流量数据要求消除假日效应后，所得到的序列的年度总和大致等于调整之前的原序列的年度总和，存量数据的春节调整取决于当月月底是否在春节影响时效之中，不存在互补性约束

流量数据的春节效应-I

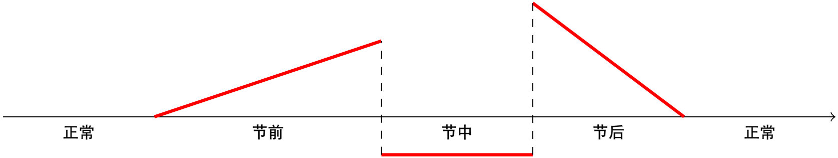
- 传统的季节性调整原理是对于给定月份 j，受到节日影响的时段落在 j 月份的时间占整个受影响时段 w 的比例记为：

$$P(w,\ j) = \frac{1}{w}\text{节前（中，后）w 天中落在 j 月份的天数}$$



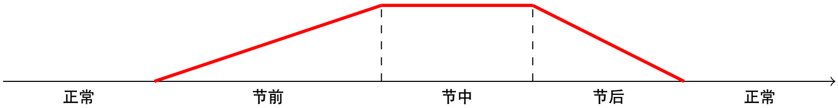
流量数据的春节效应-II

- 根据栾惠德 (2007)，改进原模型：假设节前、节中和节后春节效应影响天数分别为 w_b, w_d, w_a 天，其影响强度呈线性增长，即节前第 w_b 天的影响权数为 $\frac{1}{\sum_i^{w_b} i}$，节前第 $w_b - 1$ 天的影响权数为 $\frac{2}{\sum_i^{w_b} i}$，$\ldots$，节前第 2 天的影响权数为 $\frac{w_b-1}{\sum_i^{w_b} i}$，节前第 1 天的影响权数为 $\frac{w_b}{\sum_i^{w_b} i}$. 节日期间 w_d 天影响强度不变，节后第 1 天的影响权数为 $\frac{w_a}{\sum_i^{w_a} i}$，节后第 2 天的影响权数为 $\frac{w_a-1}{\sum_i^{w_a} i}$，$\ldots$，节后第 $w_a - 1$ 天的影响权数为 $\frac{2}{\sum_i^{w_a} i}$，节后第 w_a 天的影响权数为 $\frac{1}{\sum_i^{w_a} i}$



存量数据的春节效应

- 传统的存量春节效应就是当观测日落在影响期内，则定义变量为 1，否则为 0
- 类似地假设节前、节中和节后春节效应影响天数分别为 w_b, w_d, w_a 天，假定影响呈线性增长，节前第 w_b 天的影响权数为 $\frac{1}{w_b}$，节前第 w_b-1 天的影响权数为 $\frac{2}{w_b}$，...，节前第 2 天的影响权数为 $\frac{w_b-1}{w_b}$，节前第 1 天的影响权数为 1. 节日期间 w_d 天影响强度不变，节后第 1 天的影响权数为 1，节后第 2 天的影响权数为 $\frac{w_a-1}{w_a}$，...，节后第 w_a-1 天的影响权数为 $\frac{2}{w_a}$，节后第 w_a 天的影响权数为 $\frac{1}{w_a}$



模型选择-I

- 季节性调整主要分为两大类：加法模型和乘法模型
  - 加法模型：$Y_t = T_t + C_t + S_t + I_t, \quad A_t = T_t + C_t + I_t$
  - 乘法模型：$Y_t = T_t \times C_t \times S_t \times I_t, \quad A_t = T_t \times C_t \times I_t$

  其中 Y 是原始序列，T 是趋势因素，C 是循环因素，S 是季节因素，I 是不规则因素，A 是季调后序列
- 对于一个时间序列，如果上述四个因素相互独立，使用加法模型更好，反之使用乘法模型更好。同时如果季节性因素规模基本保持不变，不随原始序列增长而增长，则使用加法模型；如果季节因素的规模与原始序列的水平呈比例变化，则使用乘法模型。
- 在 X13-seats 包中，有自动选择模型的选项（当然也可以手动设定为加法或乘法模型）

模型选择-II

- 季节性中有两个 ARIMA 模型需要选择：非季节性 ARIMA 模型和季节性 ARIMA 模型
- ARIMA(p,d,q) 模型中有三个参数 p, d, q，分别对应模型的自回归最高阶系数 (AR)，差分系数，移动平均系数 (MA)
- 由 regARIMA 公式：

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D(y_t - \sum_i \beta_i x_{it}) = \theta(B)\Theta(B^s)a_t$$

可以选择对应的 $\phi(B), \Phi(B^s)$ 的最高阶系数，d, D 的数值以及 $\theta(B), \Theta(B^s)$ 的最高阶系数

- X13-seats 中同样也有自动选择模型的选项

参数设定

- 回归有一些无法解释的特定时间点存在异常值，需要选择异常值的种类
  - 异常值有很多种类，比如 AO，LS，TC，SO 等，具体的定义在 manual 中都有，默认有 AO 和 LS 两种异常值
- 对于每一个数据我们需要设定不同的节前，节中以及节后天数以达到最好的模型拟合和预测结果
- 由于 Variance-Bias tradeoff, 我们需要权衡 AIC，BIC 等信息准则以及异常值的个数来确定最终的模型

谢谢大家