

多变点的 AR 模型估计的功效

张学新

(中南财经政法大学统计与数学学院, 武汉 430074)

摘要: 对含有变点的不同数据生成过程, 用蒙特卡洛方法模拟研究了自回归移动平均模型分析法检测变点的功效, 实验表明, AR 模型估计对均值变点检测有较高的势。给出了利用 AR 模型检验某企业微电子标准物质质量稳定性的实例。

关键词: 自回归移动平均; 多变点检验; 检验功效; 蒙特卡罗模拟

中图分类号: C812

文献标识码: A

引言

现实经济环境下, 诸如制度变迁, 宏观经济政策等剧烈的外生冲击可能导致变量的数据生成过程发生结构突变, 若不考虑这种突变, 基于数据不变的分析判断是不准确的。如何有效的探求未知变点, 揭示规律, 寻找相应的对策统计分析尤其是经济分析的重要课题。

国内外统计学者已提出各种检验变点存在与估计的方法, 但是他们大都仅限于某个估计量的渐近分布和收敛速度的表现研究, 很少与其他检验方法作比较。Edit Gombay^[1] (2000) 比较了变点检验的 U- 统计量和顺序检验的 U- 统计量的势; Inène Gijbels, Ulku Gürker^[2] (2003) 基于机理和最小二乘法, 提出一个估计截尾变量的危险函数的单变点方法, 模拟比较研究了该估计与 Nelson-Aalen 估计、极大似然估计的性能; Chen-Pin Wang, Mahy Ghosh^[3] (2007) 基于竞争危险模型, 用贝叶斯方法诊断变点, 讨论了两种后验预测的 p 值方法的势; Konrad Nosek^[4] (2008) 模拟比较了 SIC 检验、 β - 检验、Cusum 平方检验关于线性模型是否对所有数据适合的检验的功效, 认为 Cusum 平方检验的功效最好; GARY KOOP^[5] 等 (2007) 用 Monte Carlo 模拟构造了一个 Markov 链来估计他们提出的多变点模型; Binbing Yu, Michael J^[6] 等 (2007) 用 C++ 模拟比较了 Leman (1980) 的网格搜索法和 Hudson (1966) 的加权最小二乘估计方法的功

效。国内亦很少有研究识别多变点性能的文献, 陈希孺^[7] (1991) 对一个变点的最小二乘法的性能作了小规模的模拟试验。在工矿企业, 通过有效的实验设计对测量数据进行重复性考核和稳定性考核需要比较研究不同检验方法的功效。本文先模拟研究自回归移动平均模型法识别多变点的性能, 再应用于某企业微电子标准物质生产质量的稳定性检验。

1 自回归移动平均模型变点估计

设样本 $\{y_t\}$, $t = 1, \dots, N$ 满足

$$\varepsilon_t = \begin{cases} y_t + \sum_{i=1}^m a_i y_{t-i}, & 1 \leq t \leq m \\ y_t + \sum_{i=1}^{p'} a'_i y_{t-i}, & m < t \leq N \end{cases}$$

$\varepsilon_t \sim iid(0, \sigma^2)$, 若 $a = (a_1, \dots, a_p)^\top$ 与 $a' = (a'_1, \dots, a'_{p'})^\top$ 不相等, 则称 m 为 AR(p) 模型变点。

AR(p) 模型变点的传统诊断有 LM、CUSUM、Ave-Walk 与 Exp-Wald (Andrews & Pbbberger, 1994)^[8] 等方法。在多变点情形下 LM 方法有一定的计算复杂性, 其他方法则涉及极限分布。本文用滑动窗口的 F- 检验法检测变点。

定义 $\mathcal{Y}_{n_0, n}^j = \{y_{(j-1)n_0+1}, \dots, y_{(j-1)n_0+n}\}$ 为长度为 n 移动间隔为 n_0 的第 j 个滑动窗口, $n_0 \ll n \ll N$, 我们的算法如下: (1) 用 AR(p) 模型拟合 $\mathcal{Y}_{n_0, n}^j$, 结果记为 $AR_{\mathcal{Y}_{n_0, n}^j}(p_j)$; (2) 比较, $AR_{\mathcal{Y}_{n_0, n}^j}(p_j)$, $AR_{\mathcal{Y}_{n_0, n}^{j+1}}(p_{j+1})$ 的差异, 记

收稿日期: 2010-08-09

基金项目: 中南财经政法大学研究生创新课题 (2009BLG01)

作者简介: 张学新 (1966-), 男, 湖北宜城人, 讲师, 博士, 主要从事概率论与数理统计方面的研究。

$$F^{j, j+1} = \frac{|RSS_j - RSS_{j+1}|/d}{m \ln(RSS_j / RSS_{j+1}) / (n - p_m)}$$

其中 d 是两模型参数的个数之差, $p_m = \max(p_j, p_{j+1})$, RSS_j, RSS_{j+1} 是模型的残差平方和。对预先给定的置信水平 α , 当 $F^{j, j+1} > F_\alpha(d, n - p_m)$ 时认为两模型存在显著差异; (3) 当确认 $y_{n_n}^{j+1}$ 与 $y_{n_n}^{j+1}$ 的结构有差异时, 再以集合 $S_{j_{n_n}} = \{y_k \mid |y_k - y_{(j-1)n_n}| \leq n_0\}$ 中的点 y_k 为界, $k \in I = [(j-2)n_0 + n, jn_0 + n]$, 把 $\{y_i\}$ 重新划为两个子窗 $y_{1,k}^1, y_{1,N-k}^1$ (含义见前), 再重复 (2); 关于变点存在与否, 易有:

定理 记 $F = \arg \max_{i \in I} F^{1, k+1}$, 则当 $F > C_\alpha$ 时, 认为 $\{y_i\}$ 存在变点, 其中 C_α 是与检验水平 α 对应的临界值。

变点个数 q 的确定可以参照文献 [9] 的均值变点问题的处理。在用 $AR(p)$ 拟合数据之前, 还需要定阶。本文采用 BIC 准则: 设 $BIC(i) = n \ln \frac{RSS_i}{n} + i \ln n$, 则合适的阶 $p = \arg \min_{1 \leq i \leq M} BIC(i), M \leq n - 1$ 。其中 n 是数据长度, RSS_i 是 $AR(i)$ 的拟合残差平方和。

设 R 是序列的自协方差矩阵, 向量 $r = (R_x(1), R_x(2), R_x(3), \dots, R_x(p))^T, a = (a_1, a_2, a_3, \dots, a_p)^T$, 则 Yule-Walker 程为 $Ra = r$, 我们给出基于均方差最小的参数估计的具体算法如下:

- a) $\Gamma_j = -\frac{\Delta_{j-1}}{\varepsilon_{j-1}}$
- b) $a_{j,i} = a_{j-1,i} + \Gamma_j a_{j-1,j-p}^* \quad i = 0, 1, \dots, j$
- $a_{p,i} = \begin{cases} 1 & i = 0 \\ -a_{p,i} & i = 1, \dots, p \end{cases}$
- c) $\varepsilon_j = \varepsilon_{j-1} (1 - |\Gamma_j|^2)$
- d) $\Delta_j = r_j^H a_{j-p} r_j^H = (R_x(j), R_x(j-1), \dots, R_x(1))$

初始条件: $a_0 = 1, \varepsilon_0 = R_x(0), \Delta_0 = R_x(1)$

其中的 * 表示取复数共轭运算。

2 变点检测方法的性能研究

2.1 衡量变点检测性能的方法

目前, 理论上对变点估计 (m_p, \dots, m_q) 的分布是什么尚无答案, 也就没有用精确的置信系数和置信区间估计来评价检验功效的办法。为此, 张学新^[9] (2009), 引入欧氏距离 $Ed = \|M \hat{M}\|$ 其中 $M(m_p, \dots, m_q)$ 是真实变点的位置, $\hat{M}(\hat{m}_p, \dots, \hat{m}_q)$ 是它的估计值。选取 ε 设定蒙特卡洛模拟次数 N , 计算满足 $Ed \leq \varepsilon$ 的变点估计

次数 N_{rec} 采用 $Power = \frac{N_{rec}}{N}$ 衡量变点检测的性能。

2.2 $AR(p)$ 模型检测的性能

考虑具有两个 $AR(p)$ 变点的数据生成过程:

$$y_{t+2} = \begin{cases} 0.7y_{t+1} + 0.5y_t + \varepsilon_t & t \leq 4 \\ 0.6y_{t+1} - 0.7y_t + \varepsilon_t & 4 < t \leq 7 \\ 0.2y_{t+1} - 0.5y_t + \varepsilon_t & 7 < t \leq 10 \end{cases}$$

$\varepsilon_t \sim N(0, 1)$, 初值 $y_1 = 1, y_2 = 2$ 真实变点为 $t = 7, 10$ 模拟 1000 次, 得 $AR(p)$ 模型的检测势 $Power = 90\%$, 最后一次的变点估计为 $\hat{M}(6, 11)$ 。这表明 $AR(p)$ 模型检测对 AR 生成过程的数据变点有较高的势。

再考虑分布不同的两个变点的均值突变过程, 取

$$y_t = \begin{cases} 0.10 + \varepsilon_t & t \leq 4 \\ 1.49 + \varepsilon_t & 4 < t \leq 7 \\ 2.87 + \varepsilon_t & 7 < t \leq 10 \end{cases}$$

ε_t 分别服从 $U(-0.5, 0.5), N(0, 1)$, 各模拟 1000 次, $Power = 81\%, 84\%$ 。改变各变点处的跃度, 可以观测到 $Power$ 随之增大。

最后考虑方差突变过程, 不妨取

$$y_t = \begin{cases} 0.1 + \varepsilon_t, \varepsilon_t \sim N(0, 1), t \leq 4 \\ 0.1 + \varepsilon_t, \varepsilon_t \sim N(0, 3), 4 < t \leq 7 \\ 0.1 + \varepsilon_t, \varepsilon_t \sim N(0, 5), 7 < t \leq 10 \end{cases}$$

模拟 1000 次, 得 $Power = 4\%$, 检验的功效低。

2.3 讨论

使用 $AR(p)$ 模型检测法, 设定误差项是平稳分布, 所以对均值变点检测的功效是比较高的。变点跃度越大, 说明突变越剧烈, 变点越就容易检测出; 噪声越小, 外界干扰信息少, 越有利于变点检测。本文的实验结果证实了这些理论分析。不过使用 $AR(p)$ 模型检测法有相当的计算复杂性。

3 微电子参量值均匀性与稳定性考核实例

某军工电子企业研制若干微电子标准样片的制备, 需要评定测量不确定度, 以进行微电子量值溯源。企业通过对微电子参量重复测量, 判定被测量值是否在规定的不确定度范围内, 实施均匀性考核; 通过老化试验和规定时间间隔的检验, 进行稳定性考核。为此, 企业对备检样片在规定的条件下分别作静态老化、动态老化、温度冲击试验后, 每季度进行一次测量。企业主要关心两个问题: 数据有无趋势? 数据是否随时间发生漂移?

以下数据是该企业采集样片 541245_1 的 i_{h_A} 的引脚 A8 的测量数据, 单位: μA 。- 9193.18 - 9255.49 - 9280.42 - 9305.34 - 9305.34 - 9330.27 - 9342.73 - 9348.96 - 9361.42 - 9367.65 - 8129.84 - 8179.69 - 8204.61 - 8229.53 - 8248.22 - 8266.92 - 8285.61 - 8298.07 - 8304.3 - 8310.53 - 9057.39 - 9113.48

- 9138.41, - 9157.11, - 9175.8, - 9182.03, - 9194.5,
 - 9206.96, - 9213.19, - 9231.89, - 9036.65, - 9105.21,
 - 9130.14, - 9148.84, - 9167.53, - 9180, - 9192.46,
 - 9217.39, - 9229.86, - 9248.56, 40482.2, 40525.8,
 40581.9, 40538.3, 40550.7, 40544.5, 40544.5, 40563.2,
 40581.9, 40569.4, 43863.9, 43882.6, 43895.1, 43913.8,
 43926.2, 43932.5, 43957.4, 43976.1, 43957.4, 43982.3,
 26958, 26983, 27026.6, 27039.1, 27026.6, 27039.1,
 27051.5, 27051.5, 27057.8, 27039.1, 27279.7, 27323.4,
 27155.1, 27173.8, 27111.5, 26986.8, 27130.2, 27167.5,
 27192.5, 26787.4 前 40个观测值是第一年 4个季度的,
 每季一次性观测 10个;后 40个是第二年的。

统计分析如下: 设线性回归模型 $Y_i = \beta_0 + \beta_1 T_i + \varepsilon_i$,
 $\varepsilon_i \sim i \cdot dN(0, 1)$, 检验假设 $\beta_1 = 0$ 利用 Eview s6.0软件
 对前 40个观测值检验, 结果 $F = 0.9370$, 相应 p -值
 0.3392, 表明斜率不显著, 没有观测到不稳定性。但接着
 对后 40个观测值的检验结果却是斜率显著的, 观察到趋势。
 整个 80个观测值也是可以观察到趋势的, 回归方程
 $\hat{Y} = -17551.3 + 750.49T, T = 1, 2, \dots, 80$ 但是, 修正可
 决系数 = 0.589 说明拟合效果很差。散点图也清晰的表
 明, 数据含有三个均值变点。再用 $AR(p)$ 模型对 8批数
 据变点检测, 结果是 $\hat{M}(38, 55, 64)$ 。

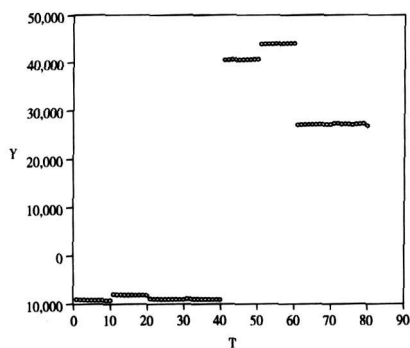


图 1 A8测量数据的散点图

在解决这个实际问题的过程中, 我们还得到一个猜想,

如果一个数据序列不能通过正态性检验, 但在某种划分下, 它的子列可以通过正态性检验, 那么, 该数据序列是适合用均值变点模型来描述的。

参考文献:

- [1] Gombay E. Comparison of U-statistics in the change-point problem and in sequential change detection[J]. Periodica Math Hungar 2000h, 41: 157-166
- [2] Irène Gijbels, Uku Gürlér. Estimation of a Change Point in a Hazard Function Based on Censored Data [J]. Lifetime Data Analysis 2003 15(9): 395-411.
- [3] Wang Chen-Pin, Malay Ghosh. Change-point diagnostics in competing risks models: Two posterior predictive p-value approaches[J]. TEST, 2007, 16: 145-171.
- [4] Konrad Nosek. Schwarz information criterion based tests for a change-point in regression models[J]. Statistical Papers, Springer Berlin/Heidelberg ISSN 0932-5026 (Print), 1613-9798 (Online).
- [5] Gay M, Koop Simon M, Potter. Forecasting and Estimating Multiple Change-Point Models with an Unknown Number of Change[J]. Federal Reserve Bank of New York Staff Reports, 2004. JEL classification: C11, C22, E17
- [6] Yu Binbing, Michael J Barretta. Estimating joinpoints in continuous time scale for multiple change-point models[J]. Computational Statistics & Data Analysis archive, 2007, 51(5): 2420-2427.
- [7] 陈希孺. 变点统计分析简介 [J]. 数理统计管理, 1991, 10(2): 52-53
- [8] Andrews D K, Pflaeger W. Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative[J]. Econometrica 1994 62(6): 1383-1438
- [9] 张学新, 段志霞. 最小二乘法对多变点检验的性能研究 [J]. 河南师范大学学报: 自然科学版, 2009, 37(6): 6-10

Efficiency of Auto regressive Moving Average Models for Change-points Detection

ZHANG Xuexin

(School of Statistics and Mathematics, Zhongnan University of Economics and Law, Wuhan 430074, China)

Abstract Based on several data generation processes with change-points, using Monte Carlo simulation, the efficiency of AR model for change-points detection was discussed firstly, and an enterprise's microelectronic standard material production quality stability's testing is given by this method secondly.

Keywords AR; changing-point test; the power of a statistical test; Monte Carlo simulation