

Applications of Nonparametric Cointegration and Error Correction Model to Finance

Jun Yin, Liyun Su*, Jiakai Zou, Xiongfei He, Hongqiang Li, Xiangwu Peng

School of Mathematics and Statistics, Chongqing University of Technology, Chongqing

Email: cloudhopping@163.com

Received: Sep. 27th, 2013; revised: Oct. 13th, 2013; accepted: Oct. 17th, 2013

Copyright © 2014 Jun Yin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. In accordance of the Creative Commons Attribution License all Copyrights © 2014 are reserved for Hans and the owner of the intellectual property Jun Yin et al. All Copyright © 2014 are guarded by law and by Hans as a guardian.

Abstract: This paper mainly focuses on co-integration theory and nonparametric method with nonlinear co-integration, which includes linear co-integration and linear error correction model, nonlinear co-integration and nonlinear error correction model, the ACE algorithm and local polynomial regression. It is clearly proved right by these analytical methods. The Matlab programming is fully exerted to realize the local polynomial regression, a nonparametric test method. In this paper, co-integration theory is clarified in details including linear theory of co-integration, linear estimation of error correction model, linear co-integration theory and tests, the nonlinear co-integration and error correction model as well as the estimation and inspection towards it. Moreover, the annotation is added for individual specifics, aiming to clarify the structures of co-integration. The existing application of time series nonlinear co-integration is put forward to serve the new method, namely the method of fusing the ridge regression nonparametric local polynomial regression. The simulation shows that this method is proved to be right. The index data assisting the researcher access to the empirical analysis are references from Japan, Singapore and Taiwan. It is on its purpose by combing the non-parametric method of local polynomial regression, co-integration and error correction model to estimate the analysis on the co-integration and error correction model. The precision of the model is assured. The local polynomial regression can be aimed to assist in explaining the significance of the non-parametric method of first derivative stock indexes in Japan, Singapore and Taiwan.

Keywords: Co-Integration Theory; The Non-Parameter; Empirical Analysis; Error Correction Model

非参数协整和误差修正模型及其在金融中的应用

殷俊, 苏理云*, 周甲凯, 何雄飞, 李泓强, 彭相武

重庆理工大学数学与统计学院, 重庆

Email: cloudhopping@163.com

收稿日期: 2013年9月27日; 修回日期: 2013年10月13日; 录用日期: 2013年10月17日

摘要: 本文主要研究了协整理论和非线性协整的非参数方法两个领域, 包括线性协整及线性误差修正模型, 非线性协整及非线性误差修正模型, ACE算法和局部多项式回归方法, 基本梳理清楚了该领域的研究脉络和框架。本文运用 Matlab 编程实现了局部多项式回归这一非参数检验方法, 详细地梳理了协整理论的内容, 包括线性协整理论、线性误差修正模型、线性协整理论的估计和检验、非线性协整和误差修正模型及其估计和检验, 并且对中细节进行了注解, 使脉络更为清晰明了, 从而增进协整理论的易读性。对时间序列协整的非线性存在的应用提出了新的方法, 即融合岭回归的局部多项式回归的非参数方法, 通过仿真表明, 该方法有很好的估计效果。选取日本、新加坡、台湾三地指数数

*通讯作者。

据进行实证分析,把局部多项式回归的非参数方法和协整、误差修正模型结合,实现了对协整、误差修正模型的估计,并且得到了较高模型估计精度,尤其重要的是,能够合理地解释局部多项式回归这一非参数方法的一阶导数在日本、新加坡、台湾三地股市指数中的意义。

关键词: 协整理论; 非参数; 实证分析; 误差修正模型

1. 引言

经济活动和金融事件的多因素性、随机波动性、事件发生的不可逆性以及时间序列的非平稳性和非线性一直都影响着经济学的科学化进程。2003年诺贝尔经济学奖授予美国经济学家 Engle 和英国经济学家 Granger,前者发明了 ARCH 模型,后者发明了对非平稳时间序列建模的协整理论。协整理论研究多个时间序列之间的长期均衡关系,但对于线性和非线性系统,它们内部的均衡关系一般不同。对于整数位时间序列,Granger 和 Engle 所建立的协整理论^[1,2]描述了系统内分量之间的长期线性均衡关系,其基本思想是通过线性组合来消除多个时间序列共同的长期趋势,并建立它们之间的线性均衡关系,因此可以称作线性协整。经过了二十多年的发展,传统的协整理论,已经形成了一套完整的体系,其理论分析也日趋成熟。但是,许多经济时间序列是分数维的,它们具有非线性和长记忆性,同时这些时间序列之间的均衡关系往往也是非线性的。因此线性协整理论并不能满足现实的需求,为了能够更有效的描述系统内部的关系,有必要应用非线性协整和误差修正建模^[3,4]方法。众所周知,协整理论和误差修正模型的核心问题就是估计和检验问题。传统的协整和误差修正程序和检验方法可能在非线性协整场合下不合适。尤其是估计程序,传统的估计程序用于非参数协整和误差修正往往存在估计功效低下或不合适的问题,因此,本文利用局部多项式的非参数估计方法,实证研究表明,这种方法能极大地提高估计功效。

2. 文献综述

2.1. 线性协整理论研究现状

国外方面:Granger(1981)提出了协整的概念,阐述了协整与误差修正模型(ECM)之间的关系,并指出了协整关系刻画了经济系统的长期均衡关系。Granger 和 Weiss (1983)进一步的提出了 Granger 表示定理,建

立了误差修正模型(ECM)形式。Engle 和 Granger (1987)正式提出协整理论,该理论为两个或者多个非平稳变量之间寻找均衡关系,以及建立误差修正模型(ECM)奠定了基础。而 ECM 的概念最初由 Sargan (1964)提出^[5],后来 Henry-Anderson(1977)^[6]和 Davidson(1987)^[7]等人进一步完善了这一概念。James H. Stock 和 Mark W. Watson(1988)将协整关系描述为同趋势^[8]。在协整关系存在性检验中,通常采用的方法是先确定分量时间序列的单整阶数,然后根据建立单一方程或系统方程进行协整检验,通常的单位根检验方法包括由 Dickey 和 Fuller(1979, 1981)提出的 DF 检验以及 ADF 检验^[9,10]。协整关系估计方面,协整关系表现形式不同,估计方法就有所不同,由 Engle 和 Granger(1987)提出的单方程协整 EG 两步法应用十分广泛;Johansen (1988, 1991)从协整向量误差修正模型(ECM)出发,讨论协整内部的长期关系估计的极大似然方法,这种方法是通过求解广义特征方程的方法来估计 ECM 中的参数的,同时还得到了协整关系存在性的似然比统计量。

国内方面:魏巍贤(1997)考虑当被预测的时间序列变量为非平稳过程时,由单项预测构成组合预测的条件,通过研究发现组合预测中的任何一个单项预测,与被预测变量具有协整关系是构成组合预测的重要条件^[11]。张丽杰(2006)利用协整检验和 Granger 因果关系检验等计量方法,对进口和出口与 GDP 的关系进行了检验,结果发现它们之间存在长期稳定的关系,但出口与 GDP,进口与 GDP 之间仅存在单向因果关系,即经济增长是进、出口增长的原因,而进、出口的增长并不是经济增长的原因^[12]。江孝感等(2008)提出了有波动持续的向量时间序列,若各分量为一阶单整且存在线性协整关系,则它们之间一定存在线性的协同持续的关系,并且协同持续向量就是协整向量^[13]。

2.2. 非平稳时间序列的非参数回归的研究现状

国外方面:Park和Phillips(1998)利用单证序列非线性变换的渐进结果考虑了单整序列的和密度估计

问题,并考虑了非参数自回归方法估计单证过程得到的一些结果。Babdi(2004)进一步与考虑了单整序列非线性回归的局部多项式的渐进性质。Karlsen, Hans A., Myklebust, Terje, Tjøstheim, Dag (2007), Schienle (2008), Wang, Qiyang, Phillips, Peter C. B. (2009a, 2009b, 2011)研究了非参数核回归的非线性渐近理论协整系统^[14-16]。

国内方面:张喜彬、孙青华和张世英(1999)对时间序列长记忆、非线性特点,提出向量时间序列的概念,用神经网络方法对其存在性进行了检验,并讨论了非线性协整关系检验的可行性,同时给出了其估计算法^[17]。孙青华、张喜彬和张世英(2000)运用神经网络技术,提出了一种估计和检验非线性长记忆时间序列之间协整关系的方法,并通过仿真试验说明了所提方法的使用价值^[18]。程细玉和张世英(2001)研究了向量分整序列的非线性问题,并对非线性协整的关系进行了讨论,给出了若干成果^[19]。孙青华和张世英(2002)在线性协整基础上,对非线性协整关系做了深入分析,并且给出了非线性协整的定义;将分形理论引入协整关系,从另一角度解释了非线性时间序列的协整关系和长记忆性存在的原因,拓展了协整问题的内涵^[20]。

2.3. 局部多项式在非参数估计方面的研究现状

国外方面:局部多项式拟合是用途广泛的非参数技术,拥有多种好的统计性质。这些性质的内容可参阅Fan和Gijbels(1996)^[21]。近年来,局部多项式方法作为一种非参数技术,得到了迅速地发展。Takeda, Farsiu和Milanfar (2006, 2007, 2008)等人将局部多项式的非参方法用于图像重建^[22]。

国内方面:何其祥和郑明(2003)将局部多项式运用到线性回归的参数估计中,并得到了较好的估计^[23]。Zhang等人(2009)运用局部多项式完成了在医学图像大脑磁共振图片的高分辨率重建,且得到较好的边缘和细节。Su Liyun和Li Fenglan将局部多项式的方法运用到DWT域迭代维纳滤波、多元混沌时间序列预测、深圳成分指数等问题中,并取得较好的效果^[24-26]。

3. 理论基础

3.1. 线性协整理论

一个有不确定性分量的时间序列 $\{X_t\}$,其经过 d

阶差分后,变为平稳、可逆的ARMA表现形式,而其 $d-1$ 阶差分仍为非平稳序列,则称时间序列 $\{X_t\}$ 为 d 阶单整,记作 $X_t \sim I(d)$, d 是大于零的整数。对 n 维向量时间序列 $\{X_t\}$,它的分量序列是 (d,b) 阶线性协整的,记作 $X_t \sim CI(d,b)$,如果(1) X_t 的分量,均为 $I(d)$ 序列,其中 $i=1,2,\dots,n$, d 为整数;(2)存在一个向量 $\alpha \neq 0$,使得 $Z_t = \alpha^T X_t \sim I(d-b)$;其中 $b > 0$, b 是整数。其中: $I(d)$ 表示 d 阶单整过程, α 称为协整向量。当 $n=2$ 时,如果协整关系存在,则协整向量唯一; $n > 2$ 时,协整关系不唯一,可能会有多个协整关系存在,如果存在 r 个独立的、线性的协整向量且形成的 $n \times r$ 阶矩阵 α ,则称之为协整矩阵。

3.2. 非线性协整理论

线性协整理论研究的是有整数维的时间序列列之间的均衡关系。对于非线性的金融市场来说,单个分量序列不是整数维的,而是分数维,那么序列的均衡关系就是非线性的。对于非线性系统而言,它与线性系统的均衡关系一般不同,通过简单的线性组合很难消除它们的共同趋势,此时,线性协整理论已不再适用,而非线性协整理论是解决这一问题的有效途径。 n 维向量时间序列 $X_t = (x_{1t}, x_{2t}, \dots, x_{nt})^T$, $\{X_t\}$ 的分量序列是非线性协整的,如果(1) x_{it} 是长记忆序列, $i=1,2,\dots,n$;(2)存在一个函数 $f(\cdot)$,使 $y_t = f(x_{1t}, x_{2t}, \dots, x_{nt})$ 是一个短记忆过程。这里函数 $f(\cdot)$ 是非线性的,称为非线性协整函数。

4. 基于局部多项式的非参数协整和修正误差模型及其应用

经过二十年的发展,世界的股票市场已经形成,股市对国民经济的“晴雨表”作用亦日益凸现,日本、新加坡、台湾三个区域性证券交易所在不同的点上同时反映着该地区国民经济的情况。由于这三个交易所位于不同地域,经济发展背景不同,分布其中的上市公司素质和投资者构成成分差异较大,交易规则也不同,因此这三个股票市场呈现出不尽相同的运行特征。为更好地认识三个地区间的经济联系,我们有必要研究三地的股市联系,利用局部多项式回归这一非参数方法进行向量时间序列的非线性协整研究,给出了非线性协整的建模方法,最后对日本、新加坡、台湾股

市指数数据进行了实证研究。

4.1. 数据的选取和处理

本文选用数据为1998年1月1日至2013年4月4日的日本、新加坡、台湾三地股市综合指数，即日经225指数、新加坡海峡时报指数、台湾加权指数。所用数据均来自RESSET金融研究数据库^[27]。

由于本文研究的是三项指数之间的长期关系，可以忽略某一个交易日数据对整体的影响，因此选用每月月末交易值作为一个样本点。另外，由于大部分的股市交易数据序列都存在指数型增长趋势，为保证数据可比性和消除可能存在的异方差，本文对原始数据进行标准化处理，后面的计算都使用的是标准化后的数据。

4.2. 证券指数的分析与格兰杰因果检验

4.2.1. 单位根检验

为了避免伪回归等问题出现，在检验变量间的协整关系前首先需要验证序列的平稳性。如果一个序列经过 p 次差分之后具有平稳性，则称该序列是 p 阶单位根过程，表示为 $I(p)$ 。本文利用ADF检验来分别对各原始变量序列的平稳性进行单位根检验，检验结果如表1(其中日经225指数 RB_t 、新加坡海峡时报指数 XJP_t 、台湾加权指数 TW_t)：

从表1中单位根检验的结果可以看出，金融时间序列 RB_t 、 XJP_t 和 TW_t 的ADF统计量的值均大于5%的临界值，表明时间序列 RB_t 、 XJP_t 和 TW_t 都存在单位根，即是非平稳序列。但其一阶差分序列在5%的水平下拒绝了原假设，即一阶差分后的序列是平稳的。因此可以确定时间序列 RB_t 、 XJP_t 和 TW_t 都是 $I(1)$ 的。因此需要对非平稳时间序列 RB_t 、 XJP_t 和 TW_t 进行Johansen协整检验，以考察日经225指数、新加坡海峡时报指数、台湾加权指数这三个金融时间序列之间是否存在长期稳定均衡的关系。

4.2.2. VAR模型的构建

对非平稳金融时间序列 RB_t 、 XJP_t 和 TW_t 进行协整检验之前，需要对其建立向量自回归模型，以便客观地确定相关变量的最大滞后期。因此，可以得到如下的 $Var(p)$ 模型：

$$Z_t = C + A_1 Z_{t-1} + A_2 Z_{t-2} + \dots + A_p Z_{t-p} + \varepsilon_t$$

其中， Z_t 是由 RB_t 、 XJP_t 和 TW_t 构成的三维向量， C 为常数项， p 为最大滞后期， ε_t 为随机扰动项。

向量自回归模型的估计结果对于变量滞后期 p 的选择敏感程度很高，所以在构建 Var 模型时，必须正确的选择合理的滞后期，表2给出了不同滞后期的各计量准则的数值，表2明合理的最佳滞后期为2期，因此我们建立 $Var(2)$ 模型。

4.2.3. 协整检验及结果分析

若几组时间序列均不是平稳的，但其差分形式是平稳的，并且有相同的单整阶数，那么这些变量的某些线性组合可能是平稳的。协整检验的目的就是验证一组非平稳序列的线性组合是否具有稳定的长期均衡关系。协整检验从检验的对象上可分为两种：一种是基于回归系数的协整检验，如Johansen协整检验；另一种是基于回归残差的协整检验，如CRDW

Table 1. Unit root test results
表 1. 单位根检验结果

检验序列	检验形式 (c, t, k)	ADF 值	P 值	%5 临界值	结论
RB_t	(c,t,13)	-1.6382	0.7741	-3.4346	不平稳
XJP_t	(c,t,13)	-2.2452	0.4614	-3.4346	不平稳
TW_t	(c,t,13)	-3.5677	0.0555	-3.4348	不平稳
∇RB_t	(nc,nt,13)	-12.1270	0.0000	-1.9426	平稳
∇XJP_t	(nc,nt,13)	-11.8336	0.0000	-1.9426	平稳
∇TW_t	(nc,nt,13)	-7.9345	0.0000	-1.9426	平稳

注：表中 ADF 检验的最大滞后期为 13， ∇ 表示序列的一阶差分，检验形式(c, t, k)表示(截距项，趋势项，滞后阶数)，(nc, nt, k)表示(没有截距项，没有趋势项，滞后阶数)其中滞后阶数 n 的选取是结合 AIC 准则和 DW 统计量的值综合确定的(DW 统计量的值应接近 2 否则残差可能存在自相关)。

Table 2. p step lag value of AIC and SC statistics
表 2. p 步滞后 AIC 和 SC 统计量的值

Lag	AIC	SC
0	7.5630	7.6154
1	-0.6359	-0.3256
2	-0.6953	-0.4254
3	2.5695	2.7815
4	3.2921	3.5049
5	3.8204	4.0341
6	4.1197	4.3342

(cointegration regression Durbin-Waion)、DF 检验、ADF 检验。

我们使用向量自回归模型确定了相关变量序列的最佳滞后期阶数为 2 阶, 因此可以采用 Johansen 协整检验的方法检验日经 225 指数 RB_t 、新加坡海峡时报指数 XJP_t 、台湾加权指数 TW_t 三个变量序列是否具有长期稳定的均衡关系。在单位根检验中发现, 标准化后的 RB_t 、 XJP_t 和 TW_t 序列都没有时间趋势项但有漂移项, 利用 Eviews 软件处理, 并选用 Intercept and trend in CE-no trend in VAR 处理方式, 所得结果如表 3 所示。

表 3 中, Johansen 迹检验中 52.663 大于 0.05 显著性水平下的临界值 42.9153, 说明在 0.05 的显著性水平下拒绝没有协整方程的原假设, 即至少有一个协整方程; 另外, 由于 18.506 小于 0.05 显著性水平下的临界值 25.8721, 说明在 0.05 的显著性水平下接受最多一个协整方程的假设, 即至多存在一个协整方程。综合判断, 变量 RB_t 、 XJP_t 和 TW_t 之间存在唯一的协整方程。同样, Johansen 最大特征值检验也表明变量 RB_t 、 XJP_t 和 TW_t 之间存在唯一的协整关系。估计出的协整关系式如下

$$RB_t = -2.344 XJP_t - 0.477 TW_t - 0.0390 \quad (4.1)$$

(0.2707) (0.18204) (0.0035)

4.2.4. 误差修正模型的建立

Johansen 协整检验的结果表明: 日经 225 指数 RB_t 、新加坡海峡时报指数 XJP_t 、台湾加权指数 TW_t 之间具有长期稳定的均衡关系。根据 Granger 和 Engle 提出的 Granger 表现定理, 一定存在描述有短期波动向长期均衡调整的误差修正模型, 因此我们建立如下的误差修正模型:

$$\Delta RB_t = \sum_{i=1}^k l_i \Delta RB_{t-i} + \sum_{i=0}^k m_i \Delta XJP_{t-i} + \sum_{i=0}^k n_i \Delta TW_{t-i} - \lambda ECM_{t-1} + u_t \quad (4.2)$$

经反复试验并通过 AIC 和 SC 统计量结合相关滞后期的显著性水平, 发现最佳滞后期为 1 期, 也就是说(4.2)式中的 $k=1$ 。误差修正项反映了日经 225 指数、新加坡海峡时报指数、台湾加权指数之间的关系偏离长期均衡状态时对短期变化的影响, 其系数 λ 反映了当变量间的短期波动偏离了长期均衡状态时, 将其调整到均衡状态的调整速度或调整力度; (4.2)式中, 解

释变量差分项的系数反映了各解释变量的短期波动对被解释变量短期变化的影响方向和影响程度。使用 Eviews 软件处理, 并去掉统计上不显著的解释变量(显著性水平取为 5%), 得到误差修正模型的估计式:

$$\Delta RB_t = 3.2149 \Delta XJP_t - 0.2729 \Delta TW_{t-1} - 0.0883 ECM_{t-1}$$

(0.97302) (0.98242)
[3.30408] [-4.8096]

(4.3)

从上面的估计结果可以看出, 误差修正项的系数为 0.0883, 说明当短期波动偏离长期均衡时, 误差修正项将以 0.0883 的力度作反向调整, 将非均衡状态回复到均衡状态。

4.2.5. 格兰杰因果关系检验

为进一步认识 RB_t 、 XJP_t 和 TW_t 之间的关系, 识别它们之间的信息传导方向, 下面利用 Granger 因果关系进行检验。Granger 因果检验的基本思想是: 若对于所有的 $m > 0$, 基于 (y_t, y_{t-1}, \dots) 预测 y_{t+m} , 得到的均方误差, 与基于 (y_t, y_{t-1}, \dots) 和 (x_t, x_{t-1}, \dots) 得到的均方误差相同, 则 y 不是由 x 引起的, 即 x 不是 y 的 Granger 原因; 否则, x 就是 y 的 Granger 原因。由于 Granger 因果检验是检验两个变量之间关系的检验且只能建立在平稳变量之间或者是存在协整关系的非平稳变量之间, 因此我们首先需要对三个变量检验两两之间的协整关系。经检验, 金融时间序列 RB_t 、 XJP_t 和 TW_t 两两之间都存在协整关系, 协整方程如下:

$$RB_t = -1.918 XJP_t$$

(0.2492)

$$RB_t = 2.1167 TW_t \quad (4.4)$$

(0.4528)

$$XJP_t = -0.7987 TW_t$$

(0.1138)

下面利用 Eviews6.0 进行了滞后期分别为 1~6 阶的 Granger 因果检验, 检验结果如表 4 所示。

由表 4 我们可以得到如下结论: 1) RB 是 XJP 的 Granger 因, 同样 XJP 也是 RB 的 Granger 因, 即 RB 和 XJP 之间有明显的相互影响 2)。RB 对 TW 滞后 1 阶, 2 阶影响, 滞后 3 阶以后无影响, 而 XJP 对 TW 无影响。3) RB 和 TW 相互之间没有直接的影响。

4.3. 实证分析的局部多项式回归

由于协整分析度量的是长期稳定的均衡关系, 不能得出日经 225 指数、新加坡海峡时报指数、台湾加

Table 3. Cointegration test results
表 3. 协整检验结果

协整个数	特征值	迹统计量	迹检验		最大特征值检验		
			0.05 临界值	p 值	最大统计量	0.05 临界值	p 值
$r \leq 0$	0.117	52.663	42.9153	0.004	34.1470	25.8232	0.0032
$r \leq 1$	0.071	18.506	25.8721	0.311	12.856	19.387	0.3398
$r \leq 2$	0.032	5.6501	12.5199	0.506	5.6501	12.518	0.506

Table 4. Granger causality test
表 4. Granger 因果检验

原假设 H	序号	结论
RB 不是 XJP 的 Granger 因	1	滞后 1~6 阶拒绝
XJP 不是 RB 的 Granger 因	2	滞后 1~6 阶拒绝
RB 不是 TW 的 Granger 因	3	滞后 1, 2 阶拒绝, 其余接受
TW 不是 RB 的 Granger 因	4	滞后 1~6 阶接受
XJP 不是 TW 的 Granger 因	5	滞后 1~6 阶接受
TW 不是 XJP 的 Granger 因	6	滞后 1~6 阶接受

权指数各个时期的变化情况, 因此我们采用非参数回归的局部多项式方法, 估计台湾加权指数对日经 225 指数、新加坡海峡时报指数的偏导数, 以此来考察日经 225 指数和新加坡海峡时报指数对台湾加权指数的影响。非参数回归是统计学中新近发展起来的一种回归方法, 由于限制条件少, 不会错误的估计回归函数, 因此受到极大地关注, 关于非参数方法的介绍可参见 Handle 等。局部多项式拟合是一个用途广泛的非参数技术, 拥有多种良好的统计性质, 见 Fan 和 Gijbels。Ruppert 和 Wand 将局部多项式的方法推广到了多元回归的情形, 并证明了在这种情况下仍能自动纠正边界偏倚。Fan 等证明了多元局部多项式拟合的渐近极小极大性及其最优核函数。考虑到前面的分析和局部多项式回归中自变量过多易产生“维数祸根”问题, 对日本、新加坡、台湾三地股市指数建立如下的回归函数:

$$TW_t = f(RB_{t-2}, XJP) + \varepsilon_t. \quad (4.5)$$

根据 Fan 和 Gijbels 的推荐, 本部分采用局部二次回归即 $p = 2$ 。由 Ruppert 和 Wand 可知, β 的估计为

$$\hat{\beta} = e(X_x^T W X_x)^{-1} X_x^T W Y \quad (4.6)$$

其中

$$X_x = \begin{bmatrix} 1 & (X_1 - x)^T & \text{vech}^T(X_1 - x)^T \\ \vdots & \vdots & \vdots \\ 1 & (X_n - x)^T & \text{vech}^T(X_n - x)^T \end{bmatrix},$$

$$Y = (Y_1, \dots, Y_n)^T,$$

$$W = \text{diag}(K_H(X_1 - x), \dots, K_H(X_n - x)).$$

根据前面的讨论, 这里我们的核函数选用 Epanechnikov 核, 即

$K(X) = \left\{ p(p+2)\Gamma(p/2) / (4\pi^{p/2}) \right\} (1 - \|X\|_+^2)$ 。另外, 对于带宽的选择, 我们采用上述重点讨论的搜索法, 并选取 $h_{\min} = 2h_m/n$, $h_{\max} = h_m/2$, 其中

$$h_m = \max \|X_i - X_j\|, \quad c = 1.1.$$

由于实际问题中数据一般较少, 矩阵 $X_x^T W X_x$ 容易出现奇异的情况, 因此本文采用岭回归的思想来估计回归函数与导数值, 即

$$\hat{\beta} = e(X_x^T W X_x + \lambda I)^{-1} X_x^T W Y, \quad (4.7)$$

其中 I 是单位矩阵, λ 是待选择的参数。这里, 为了尽可能的减少估计的偏差, 选用较小的参数 $\lambda = 0.001$ 。采用(4.7)估计的标准差为 $s = 0.0156$; 若采用局部线

性回归, 拟合的标准差为 $s = 0.0316$, 因此, 采用局部二次拟合的效果更好。

图 1 显示的是台湾加权指数(TW)的局部二次拟合和原始数据的走势。从图 1 中可以看出, 拟合数据与原始数据基本重合, 但是由于前半部分数据较为密集, 对图形进行放大后的处理(图 2), 可以看出调整后的数据和原始数据吻合程度较好。

图 3 是拟合的残差图, 可以看出残差不规则的分布在 0 上或者 0 轴两侧, 并且残差值很小。结合图 1 和图 3, 不难得出: 采用局部多项式方法对数据的拟合效果是非常好的。

图 4 显示的是台湾加权指数对前两期日经 225 指数, 新加坡海峡时报指数综合指数的偏导数。从图 4 可以看出, 日经 225 指数和新加坡海峡时报指数对台湾加权指数的影响呈现出较强的非线性特征。同时可以看出日经 225 指数, 新加坡海峡时报指数对台湾加权指数作用呈现明显的时期性, 大致可以分为 2 个时期。

第一时期(1998 年~2006 年), 这一时期内, 前两期日经 225 指数, 新加坡海峡时报指数对台湾加权指数的作用是同步变化的, 并且正作用和负作用交替出现, 影响作用非常明显。

第二时期(2007 年~2013 年), 这一时期新加坡海峡时报指数对台湾加权指数呈现出较强的正面影响, 而前两期的日经 225 指数对恒生台湾加权指数有负面影响。

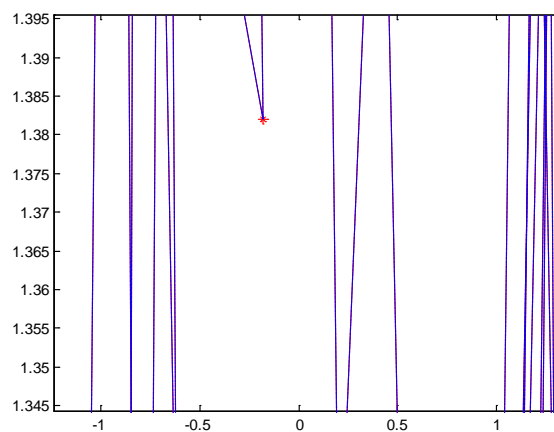


Figure 2. pictured here coordinate zero intensive part of the larger version of figure 4.1

图2. 本图为图4.1坐标0处密集部分的放大图

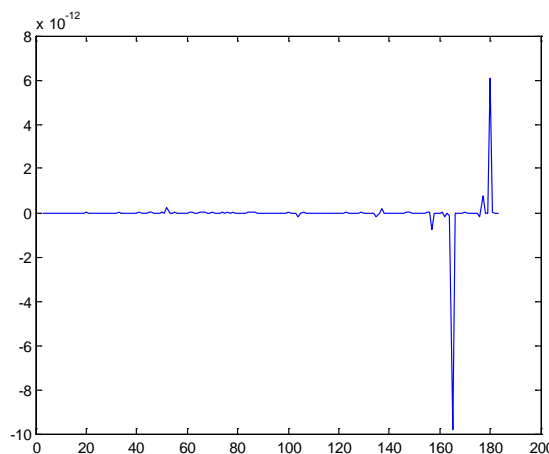


Figure 3. The weighted index of locally quadratic fitting residual figure

图3. 台湾加权指数的局部二次拟合残差图

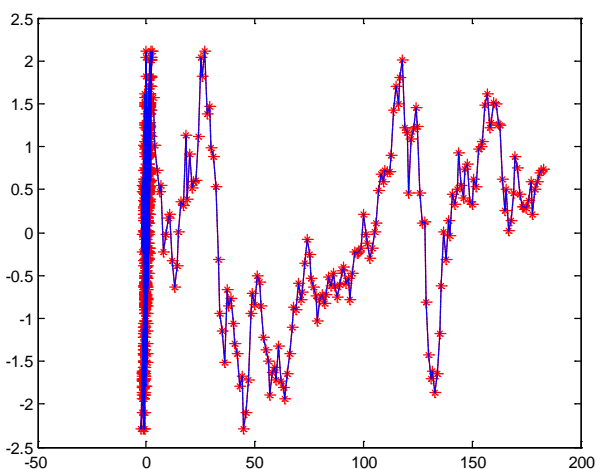


Figure 1. Fitting TAIEX index charts and the original data
图1. 拟合的台湾加权指数指数和原始数据走势图

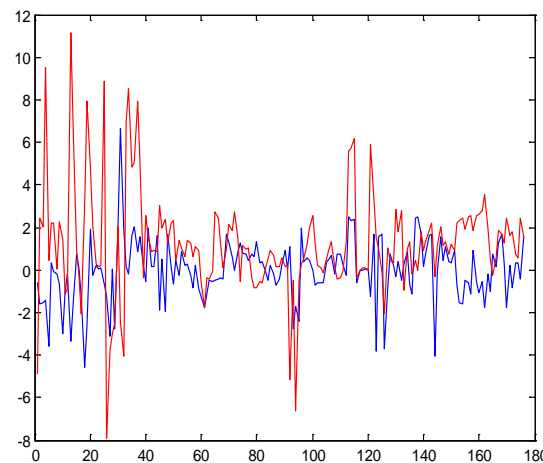


Figure 4. Partial derivative estimates
图4. 偏导数估计值

5. 总结

目前国际上对时间序列协整理论的研究及对协整关系存在性检验方法的研究还不够成熟,许多实证问题尚不能用目前已有技术解决,所以协整理论及其实证分析这一方面还有很大的研究空间。本文主要做了两方面的内容:一是对时间序列协整的非线性存在的应用提出了新的方法,即融合岭回归的局部多项式回归的非参数方法,通过仿真表明,该方法有很好的估计效果。二是选取日本、新加坡、台湾三地股市指数数据进行实证分析,把局部多项式回归的非参数方法和协整、误差修正模型结合,实现了对协整、误差修正模型的估计,并且得到了较高模型估计精度,尤其重要的是,能够合理地解释局部多项式回归这一非参数方法的一阶导数在日本、新加坡、台湾三地股市指数的中的意义。研究表明,关于协整关系自理论到实际,尚有大量的工作和问题需要完善和解决,本文只涉及了其中的一小部分,由于能力所限,可能会有错误和遗漏,权当抛砖引玉。

基金项目

重庆市教委研究生教育教学改革研究项目(Yjg133029)、重庆理工大学研究生教育教学改革研究项目(yjg2012208)、统计学校级特色专业建设项目、校级重大教学成果培育项目(统计应用能力培养的研究与实践)。

参考文献 (References)

[1] Granger, C.W.J. (1981) Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, **16**, 121-130.

[2] Engle R.F., Granger C. W. J. (1987) Cointegration and error correction: Representation, estimation, and testing. *Econometrica*, **55**, 251-276.

[3] 樊智, 张世英 (2005) 非线性协整建模研究及沪深股市实证分析. *管理科学学报*, **1**, 73-77.

[4] 刘丹红, 张世英 (2006) 基于小波神经网络的非线性误差校正模型及其预测. *控制与决策*, **10**, 1114-1118.

[5] Sargan, J.D. (1964) Wages and prices in the United Kingdom: A study in econometric methodology. In: Hart, P.E., Mills, G. and Whitacker, J.K., Eds., *Econometric Analysis for National Eco-*

nomical Planning, Butterworths, London, 34-36.

[6] Anderson H. and Berra P. (1977) Minimum cost selection of secondary indexes for formatted files. *ACM Transactions on Database Systems*, **2**, 68-90.

[7] Davidson (1987) Asymptotic properties of least squares estimators of cointegrating vectors. *Econometrica: Journal of the Econometric Society*, **55**, 1035-1056.

[8] Stock J.H. and Watson, M.W. (1988) A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica*, **61**, 783-820.

[9] Dickey, D.A. and Fuller, W.A. (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, **74**, 427-431

[10] Dickey, D.A. and Fuller, W.A. (1981) Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica*, **49**, 1057-1072.

[11] 魏巍贤 (1997) 非平稳时间序列预测组合的条件. *预测*, **4**, 47-49.

[12] 张丽杰 (2006) 辽宁省对外贸易与经济成长的协整及因果关系检验. *统计教育*, **3**, 53-55.

[13] 江孝感, 王利, 朱涛 (2008) 向量金融时间序列协整与协同持续关系——基于理论的思考. *管理工程学报*, **1**, 78-81.

[14] Karlsen, H.A., Myklebust, T. and Tjøstheim, D. (2007) Nonparametric estimation in a nonlinear cointegration model. *The Annals of Statistics*, **35**, 252-299.

[15] Schienle, M. (2008) Nonparametric nonstationary regression. Doctoral Thesis, Mannheim University, Mannheim.

[16] Wang, Q.Y. and Phillips, P.C.B. (2009) Asymptotic theory for local time density estimation and nonparametric cointegrating regression. *Econometric Theory*, **25**, 710-738.

[17] 张喜彬, 孙青华, 张世英 (1999) 非线性协整关系及其检验方法研究. *系统工程学报*, **1**, 57-68.

[18] 孙青华, 张喜彬, 张世英 (2000) 非线性协整关系的存在性研究. *管理科学学报*, **3**, 65-74.

[19] 程细玉, 张世英 (2001) 向量分整序列的非线性协整研究. *系统工程理论方法应用*, **1**, 85-87.

[20] 孙青华, 张世英 (2002) 长记忆向量时间序列的非线性协整关系研究. *天津大学学报*, **3**, 327-331.

[21] Fan, J. and Gijbels, I. (1995) Data-driven bandwidth selection in local polynomial fitting: variable bandwidth and spatial adaptation. *Journal of the Royal Statistical Society, Series B*, **57**, 371-394.

[22] Takeda, H., Farsiu, S. and Milanfar, P. (2006) Robust kernel regression for restoration and reconstruction of images from sparse noisy data. *Process of the International Conference on Image Processing*, **24**, 1257-1260.

[23] He, Q.X. and Zheng, M. (2003) Local polynomial regression for heteroscedasticity in the simple linear model. *Systems Engineering-Theory Methodology Applications*, **12**, 153-156.

[24] Su, L.Y. and Li, F.L. (2010) Deconvolution of defocused image with multivariate local polynomial regression and iterative wiener filtering in DWT domain. *Mathematical Problems in Engineering*, Article ID: 605214, 14 Pages.

[25] Su, L.Y. (2010) Prediction of multivariate chaotic time series with local polynomial fitting. *Computers & Mathematics with Applications*, **59**, 737-744.

[26] Su, L.Y. (2011) Multivariate local polynomial estimation with application to Shenzhen component index. *Discrete Dynamics in Nature and Society*, **2011**, 1-11.

[27] <http://www.resset.com/cn/>