

文章编号: 1001-0920(2006)10-1114-05

基于小波神经网络的非线性误差校正模型及其预测

刘丹红^a, 张世英^b

(天津大学 a 理学院, b 管理学院, 天津 300072)

摘要: 针对非线性系统的预测问题, 在线性和非线性协整理论涵义的基础上, 提出利用小波神经网络进行非线性协整系统的非线性误差校正模型的研究, 并给出该模型的建模方法. 对沪深股市进行实证研究, 与线性向量自回归模型进行比较. 研究证明, 小波神经网络所建立的非线性误差校正模型有较好的预测效果, 能够有效地预测非线性经济系统.

关键词: 协整; 非线性协整; 小波神经网络; 非线性误差校正模型; VAR 模型

中图分类号: F830 **文献标识码:** A

Nonlinear Error Correction Model and Forecasting Based on Wavelet Neural Networks

L I U D a n - h o n g ^a, Z H A N G S h i - y i n g ^b

(a School of Science, b School of Management, Tianjin University, Tianjin 300072, China Correspondent: L I U D a n - h o n g, E m a i l: liudanhong@eyou.com)

Abstract: The forecasting problem of the nonlinear system is discussed. Using wavelet neural network approach, nonlinear error correction model of nonlinear cointegration system is studied based on the signification of linear and nonlinear cointegration. The modeling method of nonlinear error correction model is proposed. The Shanghai and Shenzhen stock markets are analyzed. The results validate more validity of nonlinear error correction model on the wavelet neural network than linear vector autoregressive model, and forecast validly the nonlinear economy system.

Key words: Cointegration; Nonlinear cointegration; Wavelet neural network; Nonlinear error correction model; VAR model

1 引 言

文献通过协整系统的表现定理指出协整系统具有 3 种主要的表现形式, 即向量自回归 (VAR) 模型、向量移动平均 (VMA) 模型和误差校正模型 (ECM). 协整的概念是假设向量时间序列中每个分量序列都是单整的, 如果这些分量序列的某种线性组合是一个平稳序列, 那么这些分量称为协整的. 所以向量时间序列中的协整关系反映了多变量时间序列的发展趋势, 描述的是分量序列之间的长期线性均衡关系, 并且可以通过误差校正模型表现出来. 误差校正模型克服了传统时间序列建模的缺点, 它不

仅反映了水平序列的长期均衡关系, 而且可以避免伪回归现象. 基于线性协整技术的 ECM 用来提高经济预测质量方法, 使得协整技术广泛应用于经济预测中^[1-3]. 然而, 对于非线性市场, 线性协整理论对于数据的拟合与预测有着明显的局限性.

非线性协整关系普遍存在于时间序列中, 非线性协整关系使预测非线性经济系统成为可能^[4,5]. 近年来, 小波分析以其良好的时频特性成为数据处理和非线性系统辨识中的重要工具. 小波神经网络具有任意逼近非线性函数的能力, 同时由于小波变换良好的时频特性, 可以通过对尺度参数和平移参数的训练更快、更好地逼近非线性函数. 本文进一步提

收稿日期: 2005-06-27; 修回日期: 2005-10-21

基金项目: 国家自然科学基金项目 (70471050).

作者简介: 刘丹红 (1972—), 女, 吉林榆树人, 讲师, 博士, 从事数量经济的研究; 张世英 (1936—), 男, 北京人, 教授, 博士生导师, 从事系统工程、数量经济等研究.

出利用小波神经网络进行非线性误差校正模型 (NECM) 的研究, 给出非线性误差校正模型的建模方法, 对沪深股市进行实证研究, 证明了沪深股市存在非线性协整关系。与线性 VAR 模型相比, 表明小波神经网络所建立的非线性误差校正模型 (NECM) 有更好的预测效果, 能更有效地预测非线性经济系统。

2 线性协整与线性误差校正模型

定义 1^[6] (单整) 一个具有不确定性分量的时间序列 $\{X_t\}$ 经过 d 阶差分后, 具有平稳的、可逆的 ARMA 表现形式, 而该序列的 $d-1$ 阶差分仍为非平稳的, 则称序列 X_t 为 d 阶单整, 记为 $X_t \sim (d)$, d 为大于零的整数。

定义 2^[7] (协整) 对于 n 维向量时间序列 $\{X_t\}$, 若满足下列条件则称其分量序列为协整的: 1) X_t 的分量 X_{it} 均为 $I(d)$ 序列, $i = 1, 2, \dots, n$, 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: $I(d)$ 表示 d 阶单整过程, α 称为协整向量, b 为整数且 $b > 0$ 。

线性协整理论中单个分量序列是整数维单整序列, 分量之间的线性均衡关系通过协整向量 α 来刻画。

定义 3^[7] 称 N 维向量时间序列 $\{X_t\}$ 具有误差校正的表现形式, 如果它可以表示为

$$A(L)(1-L)X_t = -\gamma Z_{t-1} + \epsilon_t, \quad (1)$$

其中: ϵ_t 是 N 维正态白噪声; $A(L)$ 是关于滞后算子 L 的矩阵多项式, 且 $A(0) = I_n$, $A(L)$ 的元素有限; $Z_t = \alpha^T X_t$ 表示 X_t 的 r 个协整关系所分别产生的均衡误差 (α 是向量时间序列 X_t 的协整矩阵); $Z_t = (z_{1t}, z_{2t}, \dots, z_{rt})^T$ 是模型的误差校正项; γ 是一个非 0 的 $n \times r$ 阶参数矩阵。

ECM 的一般形式^[7] 为

$$A(L)\Delta X_t = -\gamma Z_{t-k} + \epsilon_t = -\alpha^T X_{t-k} + \epsilon_t \quad (2)$$

其中: γ 是 $n \times r$ 维的系数矩阵, α 是 $\{X_t\}$ 的协整矩阵, k 为滞后阶数, 当 $k=1$ 时为形式 (1) 的误差校正模型 (ECM)。

ECM 是序列 $\{X_t\}$ 中分量序列之间协整关系的一种表现形式, 只有当分量序列之间存在协整关系时, ECM 才有意义。在 ECM 的表达式中, $t-k$ 时刻的均衡误差作为解释变量。当 $\{X_t\}$ 的分量序列之间存在协整关系时, 如果 $\{X_t\}$ 处于非均衡状态, ECM 描述了 $\{X_t\}$ 向新的均衡点逐渐调整的过程, 能够很好地反映一个向量时间序列的状态。

协整分析的一个重要应用就是预测, 利用 ECM 可以对具有线性协整关系的经济系统作预测。因为

ECM 的线性特点, 文献 [8] 指出用 ECM 模型进行预测增加了预测误差序列与原始序列具有相同的协整结构的约束条件, 而传统的 VAR 则不然。所以, ECM 长期预测效果是以牺牲短期预测效果为代价, ECM 的长期预测效果将优于 VAR 模型的预测效果。将文献 [1] 的案例分析与 VAR 模型相比, ECM 在提高中长期预测精度的同时, 也提高了短期预测精度, 只是与 VAR 模型的差别不大。由此可见, 线性协整系统的 ECM 对不同时期的预测效果是不同的。

3 非线性协整与非线性误差校正模型

对于具有非线性特性金融市场, 单个分量序列如果不是整数维单整序列, 而是分数维或者整数维数不同, 那么此时线性协整理论不再适用, 解决这一问题的有效途径是非线性协整理论。

定义 4^[9] (非线性协整) 对于向量时间序列 $X_t = (x_{1t}, x_{2t}, \dots, x_{nt})^T$, $\{X_t\}$ 的分量序列称为非线性协整的, 若: 1) x_{it} 是长记忆序列, $i = 1, 2, \dots, n$; 2) 存在一个函数 $f(\bullet)$, 使得 $y_t = f(x_{1t}, x_{2t}, \dots, x_{nt})$ 是一个短记忆过程。其中函数 $f(\bullet)$ 一般是非线性的, 称作非线性协整函数。

非线性市场中, 时间序列之间的相关关系往往是非线性的, 此时, 线性协整向量 α 无法揭示变量之间的共同趋势, 无法建立线性的 ECM, 因此只有通过非线性协整函数 $f(\bullet)$ 来刻画系统内部的非线性均衡特性。由于非线性系统更具一般性, 非线性协整也更具一般性, 而当系统是线性时, 线性协整便成为非线性协整的一个特例。

一些学者用非线性函数来代替式 (1) 中的误差校正项, 建立 NECM。文献 [10] 对非线性函数采用三次多项式的形式, 建立英国货币需求模型, 得到较好的拟合效果和预测效果。

对于非线性协整系统, 本文采用非线性协整函数给出误差校正项。

定义 5 对于向量时间序列 $X_t = (x_{1t}, x_{2t}, \dots, x_{nt})^T$, $\{X_t\}$ 的分量序列为非线性协整关系, 则非线性协整系统 $\{X_t\}$ 有非线性误差校正模型

$$\Delta X_t = \Gamma_0 + \sum_{j=1}^{k-1} \Gamma_j \Delta X_{t-j} + \Gamma_k f(X_{t-k}; \theta) + \epsilon_t \quad (3)$$

其中: $f(\bullet)$ 为非线性协整函数, θ 为参数向量, $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}, \dots, \epsilon_{nt})^T$ 为随机误差序列, $\Gamma_j (j = 0, 1, \dots, k)$ 为 $n \times 1$ 的系数矩阵, k 为滞后阶数。则

$$\Delta x_{it} = \gamma_{i0} + \sum_{j=1}^k \gamma_{ij} \Delta x_{i(t-j)} + \gamma_{ik} f(X_{t-k}; \theta) + \epsilon_{it} \quad (4)$$

为 $\{X_t\}$ 第 i 个分量的 NECM.

将参数估计出来, 则有相应的 NECM 第 i 个分量的 h 步预测模型, 为

$$\Delta x_{i(t+h)}|_t = \hat{y}_{i0} + \sum_{j=1}^k \hat{y}_{ij} \Delta x_{i(t+h-j)} + \hat{y}_{ik} f(x_{i(t+h-k)}; \hat{\theta}), \quad i = 1, 2, \dots, n, \quad (5)$$

$$x_{i(t+h)}|_t = x_{i(t+h-1)} + \Delta x_{i(t+h)}|_t, \quad i = 1, 2, \dots, n. \quad (6)$$

NECM 中的非线性协整函数的估计 $f(x_{t-k}; \hat{\theta})$ 是系统 $\{X_t\}$ 在 $t-k$ 时刻的均衡误差, 其作用是校正 t 时刻的差分序列值, 类似于线性协整系统中的误差校正项的作用. 因为 NECM 不具有 ECM 线性特点, 也不具有预测误差序列与原始序列相同的协整结构的约束条件^[10], 随 h 的取值不同 NECM 的非线性协整关系 $f(x_{(t+h-k)}; \hat{\theta})$ 不同, 所以 NECM 对短期、中期、长期预测效果是独立的, 这点与线性协整系统中的 ECM 对预测效果不同.

4 NECM 的建模

NECM 的建模步骤如下:

Step 1: 检验分量序列不具有线性协整关系, 证明分量序列的整数维单整阶数不同或者是分整阶数, 则具有非线性协整关系.

Step 2: 非线性协整函数 $f(\bullet)$ 的估计, 是非线性协整研究中的重点和难点问题, 本文进一步引入小波神经网络, 估计和检验非线性协整函数 $f(\bullet)$.

根据非线性协整系统的结构特点, 可以构造小波神经网络模型如图 1 所示.

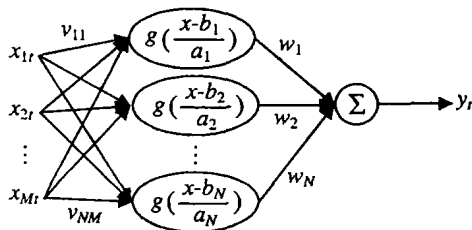


图 1 用于非线性协整建模的小波神经网络模型

网络分 3 层, 输入层 $X_t = (x_{1t}, x_{2t}, \dots, x_{Mt})^T$ 表示 M 维输入向量, $t = 1, 2, \dots, T$, T 为样本个数; 隐层节点基函数 $g(\frac{x-b_n}{a_n})$ 为小波函数, $n = 1, 2, \dots, N$; 由输入节点 j 到隐层 n 的连接权重为 v_{nj} , 由隐层 n 到输出节点的连接权重为 w_n . 需要学习的参数向量为 $\theta = (v_{11}, \dots, v_{N1}, \dots, v_{1M}, \dots, v_{NM}, a_1, b_1, \dots, a_N, b_N, w_1, \dots, w_N)^T$.

上面所示的小波神经网络可实现映射

$$f: R^N \rightarrow R, \quad (7)$$

并具有函数形式

$$y_t = \sum_{n=1}^N w_n g\left(\left(\sum_{m=1}^M v_{nm} X_{mt} - b_n\right) / a_n\right). \quad (8)$$

对于网络中参数的训练, 可以通过极小化如下形式的能量函数来实现:

$$E(\theta) = \frac{1}{2} \sum_{t=1}^T (y_t - \hat{y}_t)^2. \quad (9)$$

其中 \hat{y}_t 为教师输出, 这里取均值为常数的短记忆时间序列. 优化方法采用变尺度法进行参数训练, 直到满足结束条件或一定的收敛精度.

Step 3: 检验输出序列 L_0 ^[11] 提出了用于长记忆检验的修正 R/S 统计量 Q_n , Q_n 的分布函数为

$$F(v) = 1 - 2 \sum_{k=1}^{\infty} (1 - 4k^2 v^2) e^{-2(kv)^2}. \quad (10)$$

当长记忆不显著时, 网络输出序列 $\{y_t\}$ 为一短记忆时间序列, 说明向量 $\{X_t\}$ 存在非线性协整关系, 映射 $f(\bullet)$ 即为 $\{X_t\}$ 向量序列的非线性协整函数.

Step 4: 将得到的非线性协整函数 $f(\bullet)$ 代入 NECM (3) 中, 利用最小二乘法估计参数.

Step 5: 将参数代入预测模型 (5) 和 (6) 中, 计算结果即为预测值.

5 实证

首先选择 1999 年 01 月 04 日 ~ 2000 年 12 月 29 日的上证综合指数和深证成分指数对数收益序列共 478 对样本, 对两个样本进行长记忆性检验. Q 统计量分别为 2.140, 29.1937. 75. 可以证明在任意显著性水平下, 两个序列的长记忆性都非常显著 ($F(3) = 0.99999$), 而且分数维分别为 0.157, 0.116, 不相等, 所以, 线性协整的研究方法将不再适用. 本文选择相同数据 478 对样本, 对 2001 年 01 月 02 日 ~ 2001 年 01 月 08 日共计 5 对样本作预测, 478 对样本作一步预测, 然后增加一对样本, 再作一步预测, 共计 10 个预测值, 最后与 VAR 模型预测值作比较.

5.1 建立模型

采用 SIC 准则和 HQC 准则确定 NECM 的滞后阶数 $k = 1$, VAR 的滞后阶数 $k = 2$. 令 $\{sh_t\}$ 和 $\{sz_t\}$ 为上证综合指数和深证成分指数对数收益序列. 设

$$\{X_t\} = (sh_t, sz_t)^T, \quad t = 1, 2, \dots, n,$$

NECM 模型为

$$\Delta sh_t = \gamma_{10} + \gamma_{11} \Delta sh_{t-1} + \gamma_{12} \Delta sz_{t-1} + \gamma_{13} f(X_{t-1}; \theta) + \epsilon_{1t}, \quad (11)$$

$$\Delta sz_t = \gamma_{20} + \gamma_{21} \Delta sh_{t-1} + \gamma_{22} \Delta sz_{t-1} + \gamma_{23} f(X_{t-1}; \theta) + \epsilon_{2t}. \quad (12)$$

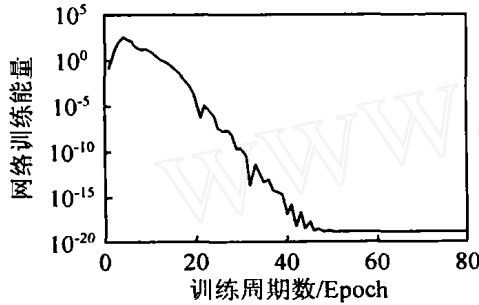
VAR 模型为

$$sh_t = \alpha_{10} + \alpha_{11}sh_{t-1} + \alpha_{12}sz_{t-1} + \alpha_{13}sh_{t-2} + \alpha_{14}sz_{t-2} + \epsilon_{1t}, \quad (13)$$

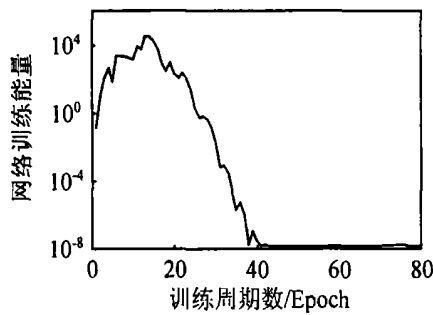
$$sz_t = \alpha_{20} + \alpha_{21}sh_{t-1} + \alpha_{22}sz_{t-1} + \alpha_{23}sh_{t-2} + \alpha_{24}sz_{t-2} + \epsilon_{2t} \quad (14)$$

5.2 模型估计

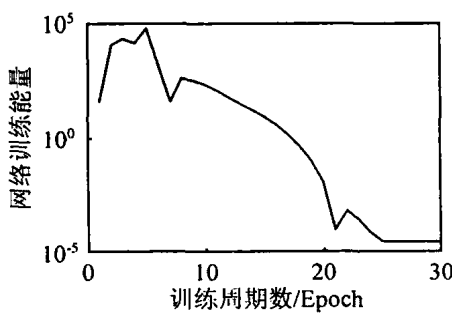
对于NECM, 首先建立小波神经网络方法拟合 5 个非线性协整函数, 建立具有 2 个输入节点, 8 个隐层节点和 1 个输出节点的单隐层小波神经网络,



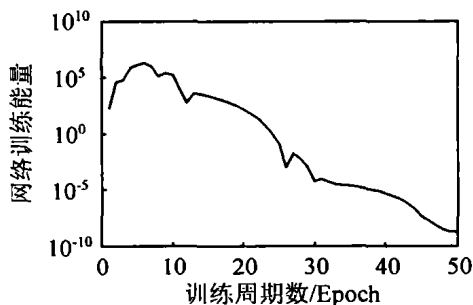
(a) 478 对样本



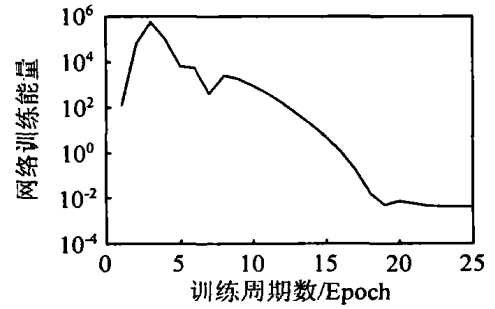
(b) 479 对样本



(c) 480 对样本



(d) 481 对样本



(e) 482 对样本

图 2 网络训练过程中能量函数

小波基函数采用了Mallet小波, 输入 5 对样本, 使用具有自适应学习速率的变尺度法对网络参数进行训练, 取学习步长为 0.1, 动量系数为 0.3 网络训练过程中能量函数见图 2

将得到的 5 组参数代入式(12)和(13)的非线性函数 $f(X_{t-1}; \theta)$ 中, 对 5 个输出序列利用修正 R/S 统计量 Q 进行长记忆检验, 表 1 给出了统计检验的结果

表 1 Q 统计检验

样本 / 对	478	479	480	481	482
Q	0.8413	0.8714	0.9175	0.8714	0.7906
$F(Q)$	0.0390	0.0563	0.0912	0.0563	0.0186

由表 1 可以看出检验结果均不显著, 证明 5 个输出序列为短记忆序列, 则该非线性函数为非线性协整函数

将非线性协整函数代入式(12)和(13)中, 使用 Eviews 软件进行最小二乘参数估计. 对于 VAR 模型同样使用 Eviews 软件进行最小二乘参数估计.

5.3 预测

NECM 预测模型为

$$\Delta sh_{t+1|t} = \hat{\gamma}_{10} + \hat{\gamma}_{11}\Delta sh_t + \hat{\gamma}_{12}\Delta sz_t + \hat{\gamma}_{13}f(X_t; \hat{\theta}), \quad (15)$$

$$\Delta sz_{t+1|t} = \hat{\gamma}_{20} + \hat{\gamma}_{21}\Delta sh_t + \hat{\gamma}_{22}\Delta sz_t + \hat{\gamma}_{23}f(X_t; \hat{\theta}). \quad (16)$$

VAR 预测模型为

$$sh_{t+1|t} = \hat{\alpha}_{10} + \hat{\alpha}_{11}sh_t + \hat{\alpha}_{12}sz_t + \hat{\alpha}_{13}sh_{t-1} + \hat{\alpha}_{14}sz_{t-1}, \quad (17)$$

$$sz_{t+1|t} = \hat{\alpha}_{20} + \hat{\alpha}_{21}sh_t + \hat{\alpha}_{22}sz_t + \hat{\alpha}_{23}sh_{t-1} + \hat{\alpha}_{24}sz_{t-1} \quad (18)$$

其中 $t+1|t$ 表示 t 时刻对 $t+1$ 时刻的预测

最后将估计出的参数代入预测模型可分别得到 10 个预测值, 预测结果如表 2 所示

表 2 中预测结果的NECM 预测误差都比较小, 除了 10 个预测值中有 3 个预测效果不如 VAR 预

表2 预测结果

序列/对		478	479	480	481	482
实际值	sh	0.014 4	0.009 7	- 0.003 1	0.003 7	- 0.011 0
	sz	0.009 5	0.005 2	- 0.004 5	- 0.001 9	- 0.015 3
NECM	sh	0.002 7 (0.011 7)	0.011 3 (0.001 6)	0.012 2 (0.013 1)	0.003 6 (0.000 2)	- 0.000 7 (0.001 03)
	sz	0.005 9 (0.003 6)	0.008 4 (0.003 2)	0.007 6 (0.012 1)	0.001 0 (0.002 9)	- 0.004 6 (0.001 07)
VAR	sh	0.001 1 (0.013 3)	0.000 0 (0.009 6)	0.000 3 (0.003 3)	0.000 9 (0.002 8)	0.000 4 (0.011 4)
	sz	0.000 8 (0.008 6)	- 0.001 1 (0.006 3)	- 0.000 7 (0.003 8)	0.000 5 (0.002 4)	- 0.001 1 (0.014 2)

注: 数据第1行为预测值, 第2行为误差

测, 其他预测都比较好。这些预测结果与实际数据规律和本文采用的滚动预测有关, 下一个预测结果受到前一个预测结果的影响, 而不是实际值的影响。因此, 预测结果进一步说明了对于具有非线性关系的时间序列, 非线性模型可以对大多数数据进行拟合, 而线性模型只适用于个别数据, 而对非线性关系的数据无法进行拟合。

6 结 论

金融市场的非线性特性使得原有的线性协整理论不再适用, 非线性协整理论可以更好地研究非线性市场中多个变量之间的相关关系。本文在已有的非线性协整理论基础上, 对利用小波神经网络建立的NECM作了一步预测, 与线性模型相比能较好地经济系统进行短期预测。因为NECM对短期、中期、长期预测结果相互无关, 本文仅给出一步预测方法, 同时也可利用本文建立的NECM作多步预测, 与一步预测方法相同, 即对不具有线性协整关系的经济系统进行中长期的预测。

参考文献(References)

- [1] 朱辉, 张世英. 协整系统的预测——中国国民收入与消费的实证分析[J]. *预测*, 1997, 6(6): 32-36
(Zhu H, Zhang S Y. Forecasting in Cointegrated System — Experiment Analysis of Chinese People Income and Consumption [J]. *Forecasting*, 1997, 6(6): 32-36)
- [2] De Gooijer J D, Vidiella-I Anguera A. Forecasting Threshold Cointegrated Systems[J]. *Int J of Forecasting*, 2004, 20(2): 237-253
- [3] Bruce E Hansen, Byeongseon Seo. Testing for Two-

regime Threshold Cointegration in Vector Error-correction Models[J]. *J of Econometrics*, 2002, 110(2): 293-318

- [4] Clement M P, Galvao A B. A Comparison of Tests of Non-linear Cointegrated with an Application to the Predictability of US Interest Rates Using Their Term Structure[J]. *Int J of Forecasting*, 2004, 20(2): 219-236
- [5] Ender W, Siklos P L. Cointegration and Threshold Adjustment [J]. *J of Business and Economic Statistics*, 2001, 19(2): 166-176
- [6] Granger C W J. Developments in the Study of Cointegrated Economic Variables[J]. *Oxford Bulletin of Economics and Statistics*, 1986, 48(3): 213-228
- [7] Johansen S. Statistical Analysis of Cointegration Vectors[J]. *J of Economic Dynamics and Control*, 1988, 12: 231-254
- [8] Engle Yoo. Forecasting and Testing in Co-integrated System [J]. *Econometrics*, 1987, 35: 143-159
- [9] 张喜彬, 孙青华, 张世英. 非线性协整关系及其检验方法研究[J]. *系统工程学报*, 1999, 14(1): 57-68
(Zhang X B, Sun Q H, Zhang S Y. Research on Non-linear Cointegrated Relation and Testing Methods[J]. *J of Systems Engineering*, 1999, 14(1): 57-68)
- [10] Hendry D F, Ericsson J G. An Econometric Analysis of UK Money Demand in Monetary Trend in the United States and the United Kingdom by Milton Friedman and Anna J Schwartz[J]. *American Economic Review*, 1991, 81(1): 8-31.
- [11] Lo W. Long-term Memory in Stock Market Prices [J]. *Econometrica*, 1991, 59(5): 1279-1313