多元长记忆 SV 模型及其在沪深股市的应用①

苏卫东[1,2]，**张世英**[1]
(1. 广东证券博士后工作站，广州 510120；2. 天津大学管理学院，天津 300072)

摘要：对多元长记忆随机波动进行建模，并给出了相应的谱似然估计方法以及在多元随机波动模型框架下分数维协整的检验步骤. 最后用上海和深圳的数据对所给的模型与方法进行了实证检验.

关键词：多元长记忆 SV 模型；分数维协整；谱似然估计
中图分类号：F830　　**文献标识码**：A　　**文章编号**：1007 - 9807(2004)01 - 0038 - 07

0　引　言

大量的实证研究[1,2]表明：资产价格的条件波动呈现出长记忆性或大范围的持续性，Breidt 等[3]发现这种长记忆性是不能用 ARCH、GARCH、EGARCH 以及短记忆 SV 模型完整描述的，于是提出长记忆随机波动(LMSV)模型(或称分整随机波动模型，FISV 模型). LMSV 模型有着很多优点：首先，它在均方意义下是定义完好(well-defined)，从而容易建立其随机特性；其次，它与 ARFIMA 模型相对应，因此它的很多统计特性就相对应地进行研究.

由于多个金融市场之间的波动必然是相互联系相互影响的. 自然会提出这样的问题：多个金融市场波动的长记忆性是否也存在一定的关系呢？要研究这一问题，就必须将 LMSV 模型推广到多元情况，在多元 LMSV 模型下考察波动之间是否存在分数维协整关系. 波动间的分数维协整关系类似于波动的协同持续(common persistence)，它对于研究规避投资风险的持续性影响具有重要的理论意义与实践价值[4].

1　多元长记忆 SV 模型及其谱似然估计

1.1　一元长记忆 SV 模型

Taylor[5]在解释金融收益序列波动的自回归行为时提出了 SV 模型，其基本形式如下：

$$y = \exp(h_t/2)\, u_t \qquad u_t \sim \mathrm{iid}(0,1) \tag{1}$$

$$h_t = \alpha + \delta h_{t-1} + \sigma_\eta \eta_t \qquad \eta_t \sim \mathrm{iid}(0,1) \tag{2}$$

其中：y 是第 t 期的(消去均值 μ 后的)收益，h_t 是对数波动，误差过程 u_t 与 η_t 互不相关，且有零均值与单位方差，二者都是不可观测的. 这里的 h_t 被假定为 AR(1) 过程，其持续参数 δ 也不可观测，当 $|\delta| < 1$ 时，SV 模型是协方差平稳的. 参数 σ_η 度量了波动扰动的标准误差. 在大多数的实际应用中都假定 u_t 与 η_t 为正态分布，从而以 h_t 为条件的日收益也是正态分布.

式(1)两边平方并取对数，可得到如下的线性形式

$$x_t = \log y^2 = E[\log u_t^2] + h_t + (\log u_t^2 - E[\log u_t^2]) = u + h_t + \varepsilon_t$$

其中：$u = E[\log u_t^2]$，$\varepsilon_t = \log u_t^2 - E[\log u_t^2]$ 是独立同分布的，其均值为零，方差为 σ_ε^2. 如果 u_t 为正态分布，则 $E[\log u_t^2] = -1.27$，$\sigma_\varepsilon^2 = \pi^2/2$[6].

为了描述均值中的长记忆成分，Granger 和 Joyeux[7] 以及 Hosking[8] 独立地提出了分整 ARMA 过程，即 ARFIMA 过程，其形式为

$$\phi(B)(1-B)^d(y-\mu) = \vartheta(B)\varepsilon_t \tag{3}$$

其中：$\phi(B) = 1 - \sum_{j=1}^{r}\phi_j B^j$，

① 收稿日期：2001 - 09 - 05；修订日期：2003 - 11 - 10.
基金项目：国家自然科学基金资助项目(70171001).
作者简介：苏卫东(1970 —)，男，山东商河人，博士生.

$$(1-B)^d=\sum_{i=1}^{\infty}\frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}B^i,$$

$$\vartheta(B)=1+\sum_{j=1}^{p}\vartheta_j B^j,|d|<0.5,\varepsilon_t \text{ 是白噪声}.$$

为了描述资产价格的条件波动呈现出的长记忆性或大范围的持续性. Breidt 等[3] 提出了一种长记忆随机波动(LMSV) 模型, LMSV 模型就是把 ARFIMA 过程纳入标准 SV 模型的框架, 其一般形式为

$$(1-B)^d\phi(B)h_t=\theta(B)\eta_t\quad \eta_t\sim \text{iid } N(0,\sigma_\eta)\tag{4}$$

最常用的长记忆 SV 模型是一阶自回归长记忆随机波动(ARLMSV) 模型

$$(1-\phi B)(1-B)^d h_t=\eta_t\quad \eta_t\sim \text{iid } N(0,\sigma_\eta)\tag{5}$$

一元 LMSV 模型的估计一般采用谱似然估计(也称频域 QML 估计). 谱似然估计是通过最大化模型的对数谱似然函数:

$$L_n(\beta)=-2\pi n^{-1}\sum_{k=1}^{[n/2]}\left\{\log f_\beta(\omega_k)+\frac{I_n(\omega_k)}{f_\beta(\omega_k)}\right\}\tag{6}$$

其中[·] 表示整数部分. 利用式(6) 得到模型的参数 $\beta=(d,\sigma_\eta^2,\sigma_\varepsilon^2,\phi_1,\cdots,\phi_p,\theta_1,\cdots,\theta_q)^{\mathrm T}$ 的估计. $\omega_k=2\pi kn^{-1}$ 是第 k 个 Fourier 频率, $f_\beta(\lambda)$ 是 $x_t=\log(y_t^2)$ 的谱密度函数, 它为

$$f_\beta(\lambda)=\frac{\sigma_\eta^2|\theta(\mathrm e^{-\mathrm i\lambda})|^2}{2\pi|1-\mathrm e^{-\mathrm i\lambda}|^{2d}|\phi(\mathrm e^{-\mathrm i\lambda})|^2}+\frac{\sigma_\varepsilon^2}{2\pi},$$
$$-\pi\le\lambda\le\pi\tag{7}$$

且

$$I_n(\omega_k)=\frac{1}{2\pi n}\left(\sum_{t=1}^{n}x_t\cos\omega_k t\right)^2+\frac{1}{2\pi n}\left(\sum_{t=1}^{n}x_t\sin\omega_k t\right)^2$$

是 k 个正规化周期图.

1.2　多元长记忆 SV 模型及其估计

1.2.1　多元长记忆 SV 模型

随机波动模型可以推广到多元情形. Harvey 等[6] 研究了多元随机波动模型的建模问题.

令 $\mathrm Y_t$ 是 $N\times1$ 向量, 其元素表示为

$$y_{it}=u_{it}(\exp\{h_{it}\})^{1/2}$$
$$i=1,2,\cdots,N;t=1,2,\cdots,T\tag{8}$$

其中: y_{it} 是序列 i 在 t 时期的观测值, $\mathrm u_t=(u_{1t},u_{2t},\cdots,u_{Nt})^{\mathrm T}$ 是多元正态向量, 其均值为 0, 协方差矩阵为 Σ_u —主对角线上的元素为 1, 非对角线上的元素为 ρ_{ij}. 和一元模型类似, 方差分量 $\exp(h_{it})$ 的对数 h_{it} 由 AR(1) 过程产生

$$h_{it}=\gamma_i+\phi_i h_{it-1}+\eta_{it}\quad i=1,2,\cdots,N\tag{9}$$

其中: $\eta_t=(\eta_{1t},\eta_{2t},\cdots,\eta_{Nt})^{\mathrm T}$ 是多元正态向量, 其均值为 0, 协方差矩阵为 Σ_η.

为了进行进一步的研究, 将式(8) 两边平方, 再取对数可以得到如下的线性形式

$$\mathrm x_t=-1.27l+h_t+\xi_t\tag{10}$$

其中: $\mathrm x_t$ 与 ξ_t 都是 $N\times1$ 向量, 其元素分别为 $x_{it}=\log(y_{it}^2)$, $\xi_{it}=\log(u_{it}^2)+1.27$, $i=1,2,\cdots,N$, l 是由 1 组成的 $N\times1$ 向量. 可以证明[6] : ξ_t 的协方差矩阵 Σ_ξ 的第 i 行 j 列个元素为 $(\pi^2/2)\rho_{ij}^*$, 其中的 $\rho_{ii}^*=1$, 且

$$\rho_{ij}^*=\frac{2}{\pi^2}\sum_{n=1}^{\infty}\frac{(n-1)!}{(1/2)_n n}\rho_{ij}^{2n},$$
$$i\ne j;\ i,j=1,\cdots,N\tag{11}$$

式中: $(x)_n=x(x-1)\cdots(x+n-1)$.

模型(9) 可进一步推广, 使 $N\times1$ 向量 $\mathrm h_t$ 是多元 AR(1) 过程, 甚至是多元 ARMA 过程

$$\Phi(B)\mathrm h_t=\Theta(B)\eta_t\quad \eta_t\sim\text{iid } N(0,\Sigma_\eta)\tag{12}$$

其中: $\Phi(B)(\Phi_0=I)$, $\Theta(B)(\Theta_0=I)$ 分别为 p 阶与 q 阶滞后矩阵多项式. 当然, 在应用中这种推广一般是没有必要的.

在式(12) 中引入分数差分算子矩阵

$$\Phi(B)\mathrm{diag}((1-B)^{d_1},\cdots,(1-B)^{d_N})\mathrm h_t=\Theta(B)\eta_t\tag{13}$$

其中: d_1, d_2, $\cdots$, d_N 是分整阶数, 其取值范围都是(- 0.5, 0.5). 这样式(8) (或式(10)) 和式(13) 就组成多元长记忆随机波动模型. 不失一般性, 本文以二元 LMSV 模型为例进行研究. 对于多元 LMSV 模型及其估计, 文献中尚未见到, 下面给出相应的模型及其估计方法. 由于此文的估计方法是建立在谱似然函数的基础之上, 首先导出多元 LMSV 模型的谱密度.

1.2.2　多元长记忆 SV 模型的谱密度

首先给出谱展式定理:

引理(谱展式定理) 如果$\{x_t\}$ $t=0,\pm1,\pm2,\cdots$为零均值平稳过程, 则在区间[- π/2, π/2] 存在一正交过程$\{Z(\lambda)\}$, 使得在所有整数 t 处有

$$x_t=\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\mathrm e^{\mathrm i\lambda t}\mathrm dZ(\lambda)$$

其中: $Z(\lambda)$ 满足如下的条件

(a) $E[\mathrm dZ(\lambda)]=0$, $-\pi/2\le\lambda\le\pi/2$;

(b) $E[|\mathrm{d}Z(\lambda)|^2] = \mathrm{d}H(\lambda)$,

$-\pi/2 \le \lambda \le \pi/2$;

这里 $H(\lambda)$ 是 x_t 的非正规化积分谱.

(c) 在不同的频率 λ_1 和 $\lambda_2(\lambda_1 \neq \lambda_2)$ 处

$$\mathrm{COV}[\mathrm{d}Z(\lambda_1), \mathrm{d}Z(\lambda_2)] = E[\mathrm{d}Z^*(\lambda_1)\mathrm{d}Z(\lambda_2)] = 0$$

定理 向量 $X_t = [x_{1t}\ x_{2t}\ \cdots\ x_{pt}]^{\mathrm{T}}$ 与 $Y_t = [y_{1t}\ y_{2t}\ \cdots\ y_{qt}]^{\mathrm{T}}$ 存在如下的关系

$$X_t = \sum_{u=-\infty}^{\infty} g(u) Y_{t-u} + \varepsilon_t \tag{14}$$

其中:$\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}, \cdots, \varepsilon_{pt}]^{\mathrm{T}}$ 是多元白噪声,y_{it} 与 ε_{jt} 互不相关;$g(u) = [g_{ij}(u)]$,$u = 0, \pm1, \pm2, \cdots$ 为这一系统的格林函数. $\psi_{ij}(\lambda) = \frac{1}{2\pi}\sum_{u=-\infty}^{\infty} g_{ij}(u)\mathrm{e}^{-\mathrm{i}\lambda u}$ 是第 j 个输入与第 i 个输出之间的传递函数,令 $\Psi(\lambda) = (\psi_{ij}(\lambda))$,则 X_t 的谱密度与 Y_t 的谱密度之间存在如下的关系

$$f_X(\lambda) = \Psi(\omega) f_Y(\lambda) \Psi^*(\omega) + f_\varepsilon(\lambda) \tag{15}$$

其中:$\omega = \mathrm{e}^{-\mathrm{i}\lambda}$,$f_\varepsilon(\lambda)$ 表示 ε_t 的谱密度,$\Psi^*(\omega)$ 是 $\Psi(\omega)$ 的共轭转置.

证 由引理

$$x_{i,t} = \int_{-\pi/2}^{\pi/2} \mathrm{e}^{\mathrm{i}\lambda t}\mathrm{d}Z_i^{(x)}(\lambda), \quad i = 1,2,\cdots,p \tag{16}$$

$$y_{j,t} = \int_{-\pi/2}^{\pi/2} \mathrm{e}^{\mathrm{i}\lambda t}\mathrm{d}Z_j^{(y)}(\lambda), \quad j = 1,2,\cdots,q \tag{17}$$

以上两式结合等式(15),易得

$$\begin{aligned}\mathrm{d}Z_i^{(x)}(\lambda) = {} & \psi_{i1}(\lambda)\mathrm{d}Z_1^{(y)}(\lambda) + \\ & \psi_{i2}(\lambda)\mathrm{d}Z_2^{(y)}(\lambda) + \cdots + \\ & \psi_{iq}(\lambda)\mathrm{d}Z_q^{(y)}(\lambda) + \mathrm{d}Z_i^{(\varepsilon)}(\lambda) \\ & i = 1,2,\cdots,p\end{aligned}$$

写为矩阵形式

$$\mathrm{d}Z^{(x)}(\lambda) = \Psi(\lambda)\mathrm{d}Z^{(y)}(\lambda) + \mathrm{d}Z^{(\varepsilon)}(\lambda) \tag{18}$$

两边乘以各自的共轭转置,并取数学期望,考虑到 $\mathrm{d}Z^{(y)}(\lambda)$ 和 $\mathrm{d}Z^{(\varepsilon)}(\lambda)$ 无关,可得

$$\begin{aligned}E[\mathrm{d}Z^{(x)}(\lambda)\mathrm{d}Z^{(x)*}(\lambda)] = {} & \Psi(\lambda) E[\mathrm{d}Z^{(y)}(\lambda)\mathrm{d}Z^{(y)*}(\lambda)]\Psi^*(\lambda) + \\ & E[\mathrm{d}Z^{(\varepsilon)}(\lambda)\mathrm{d}Z^{(\varepsilon)*}(\lambda)]\end{aligned}$$

因此有

$$f_X(\lambda) = \Psi(\omega) f_Y(\lambda) \Psi^*(\omega) + f_\varepsilon(\lambda)$$

成立. 证毕.

由此定理可以直接得到满足式(10)与(13)的序列 x_t 的谱密度为

$$\begin{aligned}f_x(\lambda) = {} & \frac{1}{2\pi} D(\omega)^{-1}[\Phi(\omega)^{-1}\Theta(\omega)] \cdot \\ & \Sigma_\eta [\Theta(\omega^{-1})\Phi(\omega^{-1})^{-1}] D(\omega^{-1})^{-1} + \frac{\Sigma_\xi}{2\pi}\end{aligned} \tag{19}$$

1.2.3 多元 LMSV 模型的谱似然估计

一元 LMSV 模型可以采用谱似然函数方法;而对多元序列 $x_t = (x_{1t}, x_{2t}, \cdots, x_{mt})^{\mathrm{T}}$ 来说,由于其谱密度是一个矩阵,需要对谱似然方法作必要的推广,构造出如下的谱似然函数:

$$\begin{aligned}L_n(\beta) = {} & -2\pi n^{-1}\sum_{k=1}^{[n/2]} \mathrm{vech}(M(\omega_k))^{\mathrm{T}} \cdot \\ & \mathrm{vech}(M(\omega_k))\end{aligned} \tag{20}$$

其中:vech 为半向量算子,它将一个 $N \times N$ 矩阵的下三角元素映射成一个 $N(N+1)/2$ 维的列向量,$M = (m_{ij})$ 是一个 m 元矩阵,其元素为

$$m_{ij} = \log|f_{ij}(\omega_k)| + I_{ij}(\omega_k)/|f_{ij}(\omega_k)|$$

$$i, j = 1,2,\cdots,m$$

这里的 $f_{ij}(\cdot)$ 是向量 x_t 的谱密度矩阵 $f_x(\cdot)$ 的元素,且

$$\begin{aligned}I_{ij}(\omega_k) = {} & \frac{1}{2\pi n}\left(\sum_{t=1}^{n} |x_{it}x_{jt}|^{0.5}\cos\omega_k t\right)^2 + \\ & \frac{1}{2\pi n}\left(\sum_{t=1}^{n} |x_{it}x_{jt}|^{0.5}\sin\omega_k t\right)^2 \\ & i, j = 1,2,\cdots,m\end{aligned}$$

对式(20)进行最大化就可以求出相应的参数,优化方法采用的是禁忌遗传算法. Breidt 等[3] 证明了一元 LMSV 模型谱似然估计的一致性,由于多元情况下模型的谱似然函数形式的复杂性,尚无法给出其一致性的严格证明. 但是,Monte Carlo 试验(由于篇幅所限,这里没有给出)的结果表明多元 LMSV 模型的这种估计方法还是较为理想的,尤其是对于长记忆参数 d 的估计.

2 多元LMSV模型框架下的分数维协整

2.1 协整与分数维协整

协整的概念涉及到两个或两个以上在短期内可能以相对不同方式进行运动的序列间的长期均

衡关系,它反映了在社会经济系统存在长期均衡这一事实.这种潜在的均衡或有规律的关系反映在所观察到的时间序列上就是它们之间的协整关系.协整的概念最初是由 Granger[9] 与 Engle, Granger[10] 提出,由 Johansen[11] 等学者发展起来的.

如果一个序列 $\{y\}$ 经过 d 次差分可以表示成为一平稳可逆 ARMA 过程,则称它是 d 阶单整的,记为 $y \sim I(d)$.当 d 为分数时,则称序列为分整的.

对于 m 维向量时间序列 $\{X_t\}$,其分量序列称为 (d, b) 阶线性协整(linear cointegration) 的,记为 $X_t \sim CI(d, b)$,如果

(1) $\{X_t\}$ 的分量序列是 $I(d)$ 的;

(2) 存在一个向量 $\alpha \neq 0$,使得 $\alpha^{T} X_t \sim I(d - b)$ $(b > 0)$,其中,α 为协整向量(cointegrated vector).

协整的研究一般都局限于 I(1)/I(0) 形式,即设定 $d = 1$ 与 $b = 1$.其一般的做法是首先对所研究的序列进行单位根检验,如果单位根存在,然后实施协整检验.协整检验一般有两种方法:Johnson 方法[11] 与 EG 两步法[10].但是在特定的背景下,特别是对于呈现出长记忆特性的高频金融时间序列来说,I(1)/I(0) 形式有着太大的局限性.如果允许协整定义中的 b,甚至是 d 为分数,则称此时的协整关系为分数维协整(fractional cointegration).

Cheung 和 Lai[12]、Baillie 和 Bollerslev[13] 研究了分数维协整的检验问题.他们首先假设 $d = 1$,然后对序列线性组合的分整阶数进行检验,结果发现所研究的序列存在分数维协整的关系.但 $d = 1$ 这一假设太强,Dueker 与 Startz[14] 指出"在进行协整检验时,原序列单(分) 整阶数的不确定性甚至比协整向量单(分) 整阶数的不确定性更为重要."他们用 US 债券收益率 $\{y_t\}$ 及 US 债券收益率 $\{y_t\}$ 与加拿大债券收益率 $\{y_t\}$ 的线性组合来估计 ARFIMA 模型,得到 $\{y_t\}$ 的分整阶数 d_1 与协整向量的分整阶数 $d' = d_1 - b$ 的估计.然而他们并没有检验 $\{y_t\}$ 的分整阶数 d_1 与 $\{y_t\}$ 的分整阶数 d_2 是否相等(两个任意的序列 $\{y_t\}$ 与 $\{y_t\}$ 不总是有相同的分整阶数),这是线性分数维协整的必要条件[15].因而,他们采用的方法仅在两序列有相同的分整阶数的前提下是有效的.

2.2　多元 LMSV 模型框架下的分数维协整

两序列的分数维协整只有在它们有共同的分整阶数(即二元 LMSV 模型中 $d_1 = d_2 = d$) 的时候才可能存在.对于 $d_1 = d_2$ 的检验,目前尚无严格的检验方法,但可以从两方面考虑:一方面用一元模型得到 d_1 与 d_2 的估计值,观察二者有无大的差异;另一方面根据似然比检验的思想,分别在限定 $d_1 = d_2$ 与不限定的两种情况下对模型(13) (在二元情况) 进行估计,比较所得到的谱似然函数值,考察 $d_1 = d_2$ 是否成立.

如果 h_{1t} 和 h_{2t} 有着相同的分整阶数,当它们存在分数维协整关系时,则存在常数 β 使得 $h_{1t} + \beta h_{2t} \sim I(d - b)$ (其中,d 为 h_{1t} 和 h_{2t} 的分整阶数),因此在二元情况下,式(13) 可以写为

$$\Phi(B)\begin{pmatrix} (1 - B)^{d} & 0 \\ 0 & (1 - B)^{d-b} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & \beta \end{pmatrix} h_t = \Theta(B)\eta_t \tag{21}$$

通过估计此模型,得到各个参数的估计值及其渐近标准误差,如果 b 与 β 的估计值都显著地异于零,则可以说两个序列的波动是分数维协整的.

通过上面的讨论,提出一种二元 LMSV 模型下分数维协整的估计与检验程序.其步骤如下:

1) 用一元方法估计分整阶数 d_1 和 d_2.这一步可以以下面两种方式进行:(1) 对数平方收益和绝对值收益的一元 ARFIMA 模型;(2) 一元 SV 模型;

2) 如果两个随机过程是分数维的,在二元 LMSV 下,检验 $d_1 = d_2$.

如果 $d_1 \neq d_2$ (此时一元估计得到的 d_1、d_2 必然相差很大),根据文献[15]中线性协整条件,h_{1t} 和 h_{2t} 不存在协整关系,就没有必要进一步讨论;如果 $d_1 = d_2$,则分别用两序列的对数平方收益之差和绝对值收益之差估计 ARFIMA 模型,如果得到的 d 的估计明显小于每个序列的相应值,则说明两者是分数维协整的,且协整向量为 $\alpha = [1\ \beta] = [1\ -1]$.

3) 如果得到分数维协整的证据,且有 $\beta = -1$(如果 β 取其它值,则 β 取相应值),估计由式(8) 和(21) 组成的二元 LMSV 模型.

3　实证检验

3.1　样本选取

上海股市始于 1990 年 12 年 18 日,深圳股市则开盘于 1991 年 9 月 19 日,但中国股市的规范始于 1993 年,因此,这里的样本区间为 1994 年 1 月 1 日到 2000

年 12 月 31 日. 由于所研究的是两股市之间的均衡关系,因此删去这期间深沪股市不同时开市的那些交易日. 本文从飞天系统得到所选区间的上海综合指数与深圳成分指数的日收盘数据,用公式

$$R_t = \ln(P_{t+1}) - \ln(P_t)$$

计算其相应时间的日收益序列,共得到 1 709 个数据;其中, P_t 是上证综指(深圳成指) 第 t 个交易日的收盘价, R_t 是其相应的日收益.

3.2 一元模型的估计结果

分析的第一个阶段是检验上证指数与深圳成指的波动过程是否呈现出长记忆性. 我们首先对波动的两个替代:对数平方收益与绝对值收益来估计一元 ARFIMA(r , d , p) 模型. 这些估计的一致性依赖于多项式滞后长度选择的有效性. Lobato[16] 通过 Monte Carlo 分析说明了长记忆模型的估计对自回归多项式的误设极为敏感. 模型的选择采用了 AIC 与 BIC 准则, 二者都选择了 ARFIMA(1 , d ,1) .

估计的结果见表 1 的第 1 栏与第 2 栏. 长记忆参数的估计表明对数平方收益与绝对值收益都是平稳的长记忆过程. 由表 1 的第 1、2 栏可以看出,对两个市场来说,估计出的长记忆参数都是对数平方收益高于绝对收益. 对于高的滞后阶, ARFIMA 过程的理论上的自相关系数近似为 $\rho_k \approx ck^{2d-1}$ ($c > 0$) . 这就说明对于高阶的滞后项, d 的值越大,过程的自相关系数越高. 从而说明对数平方收益呈现出比绝对值收益高的自相关结构.

估计的第二个阶段是估计一元 LMSV 模型. 这里估计的是(1) 和(5) 组成的 ARLMSV 模型(估计结果见表 1 的第 3 栏) . 事实上,也考察了在式(4) 中 $\phi(B)$ 和 $\theta(B)$ 为更高滞后阶的情况,结果发现对所选的两个序列都是不必要的.

由表 1 中 1 - 3 栏的估计结果可以看出,除绝对值收益得到的分整阶数外,两个市场波动的分整参数 d 并无显著的差异(尤其是对于 ARLMSV 模型的结果) .

在这个假设有效的基础上(其进一步的检验在二元 LMSV 模型下进行) ,取 $\beta = -1$,采用 ARFIMA 方法来估计协整组合的分整阶数,栏 4 与栏 5 分别是对两市场的对数平方收益之差与绝对值收益之差得到的 ARFIMA 模型的估计,进行模型的选择检验时,所有序列选择了 ARFIMA(1 , d , 1) . 结果发现两个市场的对数平方收益之差与绝对值收益之差有着比栏 1 与栏 2 中的估计更低的分整阶数. 因此就可得出,从实证的角度来讲沪深股市波动过程是分数维协整的.

表 1 一元模型的估计

		ARFIMA		ARLMSV	ARFIMA	
		绝对值收益	对数平方收益		两市场绝对值收益差	两市场对数平方收益差
上证综指	ϕ	- 0.192 190 (0.037 714 0)	0.470 412 (0.038 381)	—	- 0.089 768 (0.035 944)	0.393 440 (0.079 083)
	ϑ	- 0.989 949 (8.26 E - 5)	- 0.839 647 (0.023 468)	—	- 0.989 994 8 (0.003 254)	- 0.573 823 (0.072 639)
	d_1	0.373 327 (0.087 679)	0.403 447 (0.104 765)	0.498 757 (0.000 747)	0.180 254 (0.083 250)	0.218 914 (0.104 866)
	ϕ	—	—	0.797 226 (0.001 213)	—	—
	σ_η^2	—	—	0.032 609 (0.000 482)	—	—
深圳成指	ϕ	- 0.170 040 (0.035 113)	0.523 935 (0.033 721)	—	—	—
	ϑ	- 0.989 994 3 (7.76 E - 5)	- 0.891 097 (0.017 438)	—	—	—
	d_2	0.261 594 (0.117 535)	0.440 102 (0.113 659)	0.480 407 (0.002 039)	—	—
	ϕ	—	—	0.802 812 (0.002 129)	—	—
	σ_η^2	—	—	0.035 351 (0.000 446)	—	—

注:表中括号中的数值对 ARFIMA 模型来说是相应估计的标准差,对 ARLMSV 模型来说是相应估计的渐近标准差;所有的估计在 5 % 的显著水平下是显著的,大多数估计在 1 % 的显著水平下是显著的.

3.3 二元模型的估计结果

一元 LMSV 模型的估计结果表明,两个市场的波动是有着共同分整参数 d. 但要想进一步检验这一假设,就必须转向二元框架.

采用上面的估计方法对(8) 和(13) 组成的二元 LMSV 模型进行估计. 估计过程中,由于 SV 模

型中含有随机波动因子,至今未有如 AIC、BIC 等准则对它们进行定阶,因此,选用的模型为与一阶 ARLMSV 相对应的二元一阶自回归 LMSV 模型

$$\left[\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} - \begin{pmatrix}\phi_{11} & \phi_{12}\\ \phi_{21} & \phi_{22}\end{pmatrix} L\right] \cdot \begin{pmatrix}(1-L)^{d_1} & 0\\ 0 & (1-L)^{d_2}\end{pmatrix}\begin{pmatrix}h_{1t}\\ h_{2t}\end{pmatrix} = \boldsymbol{\eta}_t \quad (22)$$

其中: $\boldsymbol{\eta}_t$ 的协方差矩阵为

$$\Sigma_\eta = \begin{pmatrix}\sigma_{11} & \sigma_{12}\\ \sigma_{21} & \sigma_{22}\end{pmatrix}$$

模型估计的结果见表 2 的第 1 栏. 当然,也尝试了其它不同的设定,但没有发现估计出的分整阶数有显著的差异. 第 2 栏是假定分整阶数相同时,二元 ARLMSV 的估计结果. 用似然比检验思想来检验分整阶数相同这一假设,结果没有发现它们的谱似然函数值有显著的差异,因此可以说两个市场波动有着共同的分整阶数.

由表 2 的第 1、2 栏可以看出,ϕ_{12} 显著地异于零,这一系数恰好反映了深圳股市波动($h_{2,t-1}$)对上海股市波动($h_{1,t}$)会产生影响;而联系于 $h_{1,t-1}$ 与 $h_{2,t-1}$ 的系数 ϕ_{21} 不显著地异于零,表明深圳股市的波动受上海股市波动的影响不大. 因此,一定的信息源引起深圳股市的波动,这种波动然后再传递到上海股市;或者说,相对于深圳股市,上海股市对波动源的反映是滞后的.

由于两个市场的波动有着共同的分整阶数,用相同的估计方法来估计 (8) 和 (21) 组成的模型,其结果见表 2 的第 3 栏.

表 2　二元 LMSV 模型的估计

	模型(13)	模型(13) $d_1 = d_2$	模型(21)
d_1	0.482 878(0.000 003)	0.428 827(0.004 224)	0.405 445(0.000 200)
d_2	0.466 503(0.000 005)	—	—
b	—	—	0.247 810(0.000 031)
ϕ_{11}	0.369 003(0.000 031)	0.324 680(0.004 743)	0.536 064(0.000 480)
ϕ_{12}	- 0.875 517(0.000 318)	- 0.794 398(0.013 824)	- 0.013 331(0.000 118)
ϕ_{21}	- 0.008 606(0.000 061)	- 0.017 701(0.003 475)	- 0.976 619(0.000 163)
ϕ_{22}	0.425 640(0.000 079)	0.508 816(0.005 232)	0.473 397(0.000 446)
σ_{11}	0.418 525(0.000 337)	0.494 222(0.008 082)	0.487 563(0.000 541)
σ_{12}	0.368 276(0.000 016)	0.344 794(0.007 184 8)	0.305 959(0.000 022)
σ_{22}	0.498 128(0.000 100)	0.497 870(0.004 954)	0.317 700(0.000 921)
ρ	0.985 428(0.000 002)	0.983 348(0.000 312)	0.982 493(0.000 000)

注:括号中的数是相应估计的渐近标准差,所有估计在 1 % 的显著性水平下是显著的.

波动的长记忆与波动的持续性都反映了当前波动对未来波动影响的长久性,在投资中希望能消除或减弱这种影响. 从表 2 第 3 栏可以看出,两市场波动协整组合的长记忆参数估计值比单个市场波动要小得多(减小了 0.247 81),这说明通过一定的投资组合是可以减小波动的长记忆性,从而也表明了协同持续理论在实践上是可行的.

4　结束语

本文为了研究不同市场之间的波动联系,把一元 LMSV 模型推广到多元情况,并给出了多元 LMSV 模型的谱似然方法以及多元 LMSV 模型下的分数维协整的检验程序. 并用所给的模型与方法研究了上海股市与深圳股市之间波动的联系,结果发现它们有着共同的分整阶数,两者之间的波动存在分数维协整关系. 同时,在这方面的研究有两个问题有待进一步研究,即多元 LMSV 模型谱似然估计的一致性与分整阶数相等的严格统计检验程序.

参 考 文 献:

[1]Bollerslev T, Mikkelsen H O. Modeling and pricing long memory in stock market volatility[J]. Journal of Econometrics, 1996, 73: 151—184.

[2]Ding Z, Granger C W J, Engle R F. A long memory property of stock market returns and a new model[J]. Journal of Empirical Finance, 1993, 1: 83—106.

[3]Breidt F J *et al*. The detection and estimation of long memory in stochastic volatility[J]. Journal of Econometrics, 1998, 83: 325—348.

[4]李汉东,张世英. BEKK模型的协同持续性研究[J]. 系统工程学报, 2001, 16(3): 225—231.

[5]Taylor S J. Modeling Financial Time Series[M]. Chichester: John Wiley, 1986.

[6]Harvey A *et al*. Multivariate stochastic variance models[J]. Review of Economic Studies, 1994, 61: 247—267.

[7]Granger C W J, Joyeux R. An introduction to long memory time series and fractional differencing[J]. Journal of Time Series Analysis, 1980, 1: 1—29.

[8]Hosking J R M. Fractional differencing[J]. Biometrika, 1981, 68: 165—176.

[9]Granger C W J. Developments in the study of cointegrated economic variables[J]. Oxford Bulletin of Economics and Statistics, 1986, 48: 221—238.

[10]Engle R F, Granger C W J. Co-integraiton and error correction: Representation, estimation and testing[J]. Econometrica, 1987, 55: 251—276.

[11]Johansen S. Statistical Analysis of cointegrated vectors[J]. Journal of Economic Dynamics and Control, 1988, 12: 231—254.

[12]Cheung Y W, Lai K S. A fractional cointegration analysis of purchasing power parity[J]. Journal of Business & Economics Statistics, 1993, 11: 103—112.

[13]Baillie R T, Bollerslev T. Cointegration, fractional cointegration and exchange rate dynamics[J]. Journal of Finance, 1994, 49: 737—745.

[14]Dueker M, Startz R. Maximum likelihood estimation of fractional cointegration with an application to U. S. and Canadian bond rates [J]. Journal of American Statistical Association, 1998, 93: 420—426.

[15]程细玉,张世英. 向量分整序列的协整研究[J]. 系统工程学报, 2000, 15(3): 253—257.

[16]Lobato I N. A semiparametric two step estimator in a multivariate long memory model[J]. Journal of Econometrics, 1999, 90: 129—153.

[17]柯 珂,张世英. ARCH模型的诊断分析[J]. 管理科学学报, 2001, 4(2): 12—18.

Multivariate long memory SV model and its application to Shanghai and Shenzhen stock markets

SU Wei-dong[1,2], *ZHANG Shi-ying*[2]

1. Postdoctor Center of Guandong Securities Co. LTD, Guangzhou 510120, China;

2. School of Management, Tianjin University, Tianjin 300072, China

Abstract: This paper constructs a multivariate long memory stochastic volatility (MLMSV) model and provides its spectrum likelihood estimation method, as well as the testing procedures for fractional cointegration under MLMSV frame. We also test the model and method with data of Shanghai and Shenzhen stock markets to show its effectiveness.

Key words: multivariate long memory SV model; fractional cointegration; spectrum likelihood estimation