文章编号:1003-207(2010)06-0033-09

金融市场高维波动率的扩展广义正交 GARCH 模型与参数估计方法研究

刘志东[1], 薛　莉[2]

(1. 中央财经大学, 北京　100081; 2. 南京大学, 江苏 南京　210093)

摘　要:本文对 Van der Weide (2002) 的广义正交 GARCH 模型进行扩展, 提出反映金融资产收益波动性特征, 具有"杠杆效应"的广义正交 GARCH 模型。由于这种扩展的广义正交 GARCH 模型在高维数据中面临参数估计困难, 本文从交互信息理论视角研究模型的参数估计问题, 在理论上证明基于交互信息最小化的多元 GARCH 模型参数估计与基于极大似然函数参数估计的联系和区别, 并在提出的扩展广义正交 GARCH 模型框架下, 采用不同的统计技术实现基于交互信息最小化的参数估计方法, 避免了传统极大似然函数估计需要事先正确指定标准化残差概率密度函数和高维运算困难, 计算效率较高, 使多元 GARCH 模型在高维数据中可以应用。最后, 根据全球主要金融市场的 15 种股票指数数据, 通过实证研究对建立的扩展广义正交 GARCH 模型及其参数估计方法有效性进行评价与检验。实证研究表明了本文提出的扩展广义正交 GARCH 模型与参数估计方法的优势。

关键词: 交互信息; 多元 GARCH 模型; 杠杆效应; 参数估计; 动态条件相关

中图分类号: F830　　文献标识码: A

1　引言

对单个金融资产收益的波动率和不同金融资产收益之间的条件协方差矩阵研究是金融计量经济领域的主要研究内容。可以采用多元 GARCH 模型对金融市场的相关问题进行研究。相对于一元 GARCH 模型的理论和应用研究都取得很大进展, 多元 GARCH 模型的研究还处在一个比较初期的阶段。由于涉及到对协方差矩阵建立动态模型, 从技术上看主要的困难有两个: 首先, 随着维数的增加, 需要估计大量的参数, 在现有的非线性优化技术和计算机技术下实现起来通常很困难; 其次是如何保证对参数限定条件来保证协方差矩阵的正定性。传统的多元 GARCH 模型主要包括 Bollerslev、Engle 和 Wooldridge (1988)[1] 的 VEC 模型, Engle 和 Kroner(1995)[2] 的 BEKK 模型, Bollerslev(1990)[3] 常条件相关系数(CCC)模型, Tse 和 Tsui(1999)[4]、Engle (2002)[5] 的动态条件相关系数模型(DCC), Engle、Ng 和 Rothschild (1990)[6] 的因子 GARCH 模型(Factor Garch Model)。由于各种困难, 这些模型很难应用到高维波动率建模中。刘志东(2010)[7] 对多元 GARCH 模型的相关问题进行了总结。

为此, Ding(1994)[8]、Alexander 和 Chibumba (1997)[9] 把主成分分析(PCA)应用到多元 GARCH 模型中, 提出正交 GARCH 模型(Orthogonal GARCH, 简称 O-GARCH)。O-GARCH 模型采用 PCA 技术提取主成分, 用一元 GARCH 模型表示每个主成分的条件方差, 构建时变对角矩阵, 得到条件协方差矩阵。由于完全基于无条件信息(样本协方差矩阵)估计正交矩阵, O-GARCH 模型存在可识别的问题。当数据存在弱相关时候, 模型很难识别真正的正交矩阵。O-GARCH 模型的有效性是建立在主成分是条件不相关(conditional uncorrelated)基础上的, 主成分之间的无条件不相关(unconditional uncorrelated)不能确保它们条件不相关。

O-GARCH 模型的优点在于它采用两阶段参数估计方法。通过直接用一元 GARCH 模型对原始序列的主成分建模, 能得到多元 GARCH 模型的参数。但正交矩阵的假设条件太严格。后来, Van der Weide (2002)[10] 提出广义正交 GARCH(Gen-

收稿日期: 2010-04-12; 修订日期: 2010-11-01
基金项目: 国家自然科学基金资助项目(70603034, 70971145); 教育部人文社科项目(08JC790107); 中央财经大学"211 工程"三期资助项目
作者简介: 刘志东(1973-), 男(汉族), 中央财经大学, 博士, 教授, 研究方向: 现代投资管理与资产定价、金融工程与风险管理.

eralized Orthogonal GARCH，简称 GO-GARCH）模型。该模型是 O-GARCH 模型的扩展，不需要线性映射矩阵是正交的，可以是任意的可逆矩阵。由于这种扩展，GO-GARCH 模型不再保持 O-GARCH 模型的两阶段参数估计的优点。尽管 Van der Weide (2002)[10] 最初的 GO-GARCH 模型也采用两阶段参数估计方法，但由于在第一步估计中，仅仅能识别可逆矩阵的部分参数，在第二步参数估计中，对多元极大似然函数求极值时，需要同时估计一元 GARCH 模型参数和可逆矩阵剩余的参数，这不仅使得在高维计算非常困难，而且使因子载荷矩阵的估计受到指定的一元 GARCH 模型类型的影响。对于高维系统，这种极大似然估计方法的计算速度非常低。同时，与因子模型和 O-GARCH 模型相似，Van der Weide (2002)[10] 最初的 GO-GARCH 模型也都采用主成分分析(PCA)识别基础因子，基础因子是无条件不相关的(unconditionally uncorrelated)。然而，为了确保协方差矩阵是对角矩阵，还需要假设因子必须是条件不相关的(conditionally uncorrelated)。Fan 等(2008)[11] 的研究指出，尽管主成分之间是无条件不相关的，但是它们之间不一定是条件不相关的。O-GARCH 模型和 GO-GARCH 中采用的假设会导致严重的错误，并建议采用条件不相关成分或因子(conditionally uncorrelated component, CUC)模拟多元波动性。

为了克服 Van der Weide (2002)[10] 的 GO-GARCH 模型参数估计方法的不足，Boswijk 和 van der Weide (2006)[12]、Fan 等 (2008)[11] 对 GO-GARCH 模型参数估计方法进行研究。但这些方法也都需要在高维参数空间上对某一准则函数进行求极值计算。因此，随着维数 m 的增加，每种方法都可能遇到数值计算的相关问题，如：优化算法的收敛问题，或只能找到局部极值点。Broda 和 Paolella, (2008)[13]，王明进、陈奇志 (2006)[14]，王明进 (2008)[15] 等把独立成分分析(ICA)方法应用到 Van der Weide (2002)[10] 的 GO-GARCH 模型中。许启发、张世英(2007)[16]，Zhang 和 Chan, (2009)[17] 从信息理论角度研究多元 GARCH 模型参数估计问题，通过盲源信号分离(BSS)技术得到因子 GARCH 模型。这些研究和 Van der Weide (2002)[10] 最初的 GO-GARCH 模型一样，采用最经典的一元 GARCH 模型表示每个因子的波动性，这与金融资产收益非对称方差，即波动的"杠杆效应"特征不符。尽管金融资产收益"杠杆效应"在一元 GARCH 模型中得到研究，但在多元 GARCH 模型中的研究还不多。同时，经典的独立成分分析(ICA)方法存在一定不足，导致可逆矩阵和因子的估计都存在很大的误差。另外，Zhang 和 Chan (2009)[11] 对多元 GARCH 模型的检验方法也存在不足。

鉴于以上研究存在的不足，本文对 Van der Weide (2002)[10] 的广义正交 GARCH 模型进行扩展，提出反映金融资产收益波动性特征，具有"杠杆效应"的广义正交 GARCH 模型。由于这种扩展的广义正交 GARCH 模型在高维数据中面临参数估计困难，本文从交互信息理论视角研究模型的参数估计问题，在理论上分析基于交互信息最小化的多元 GARCH 模型参数估计与基于极大似然函数参数估计的联系和区别，并在提出的扩展广义正交 GARCH 模型框架下，研究如何实现基于交互信息最小化的参数估计方法。最后，根据全球主要金融市场的 15 种股票指数数据，通过实证研究对建立的扩展广义正交 GARCH 模型及其参数估计方法有效性进行评价与检验。

2 具有"杠杆效应"的扩展广义正交 GARCH 模型

考虑 $N\times 1$ 维向量随机过程 $\{x_t\}$，$\mathfrak{J}_{t-1}$ 表示由直到 $t-1$ 时间信息产生的 σ 域流，θ 表示有限参数向量，x_t 可以表示为：

$$x_t = \mu_t(\theta) + \varepsilon_t \tag{1}$$

这里 $\mu_t(\theta)$ 是条件均值向量，$\varepsilon_t = H_t^{1/2}(\theta)\eta_t$，$H_t^{1/2}(\theta)$ 是 $N\times N$ 正定矩阵。假定 $N\times 1$ 维随机向量 η_t 满足：$E(\eta_t)=0$，$\mathrm{Var}(\eta_t)=I_N$。由于 $\mathrm{Var}(x_t\mid\mathfrak{J}_{t-1}) = \mathrm{Var}_{t-1}(x_t) = \mathrm{Var}_{t-1}(\varepsilon_t) = H_t^{1/2}\mathrm{Var}_{t-1}(\eta_t)(H_t^{1/2})' = H_t$，因此，$H_t$ 是 x_t 条件协方差矩阵。H_t 和 μ_t 都依赖于未知参数向量 θ。可以单独采用 VARMA 方法(vectorial autoregressive moving average)预测 μ_t，然后采用多元 GARCH 模型预测 H_t，并根据 $\varepsilon_t = H_t^{1/2}(\theta)\eta_t$ 预测 ε_t。

假设任何非零条件均值 $\mu_t(\theta)$ 已从 x_t 中减掉，因此，$E(\varepsilon_t\mid\mathfrak{J}_{t-1})=0$。假设 ε_t 能表示为潜在因子 $y_t = (y_{1t}, y_{2t}, \cdots, y_{mt})'$ (其中成分是独立的)的非奇异变换：

$$\varepsilon_t = Ay_t = A(H_t^y)^{1/2}\eta_t \tag{2}$$

这里 A 是 $N\times m$ 非奇异矩阵，只需满足可逆条件，不一定必须是正交矩阵。$\{\eta_t\}_{t\geq 1}$ 是向量鞅差序列，满足 $E(\eta_t\mid\mathfrak{J}_{t-1})=0$，$\mathrm{Var}(\eta_t\mid\mathfrak{J}_{t-1})=I_N$。

假设 y_t 满足: $y_t \mid \mathrm{J}_{t-1} \sim N(0, H_t^y)$，其中 $H_t^y = \mathrm{diag}(h_{1t}^y, \cdots, h_{mt}^y)$。这隐含的假设为: $E(y_{it} \mid \mathrm{J}_{t-1}) = 0$, $\mathrm{Var}(y_{it} \mid \mathrm{J}_{t-1}) = h_{it}^y$, $\mathrm{Cov}(y_{it}, y_{jt} \mid \mathrm{J}_{t-1}) = 0$, $i \neq j = 1, \cdots, m$。即 y_t 的成分是条件不相关的。如果能确保协方差的平稳性, 则有 $\varepsilon_t \mid \mathrm{J}_{t-1} \sim N(0, H_t)$。因此, ε_t 的条件协方差矩阵能被表示为:

$$H_t = \mathrm{Var}(x_t \mid \mathrm{J}_{t-1}) = AH_t^y A' \tag{3}$$

ε_t 的无条件协方差矩阵为:

$$H = \mathrm{Var}(x_t) = AA' \tag{4}$$

由于 $\varepsilon_t = Ay_t = A(H_t^y)^{1/2}\eta_t$，因子 $y_t = (y_{1t}, y_{2t}, \cdots, y_{mt})'$ 能够通过线性变换得到:

$$y_t = W\varepsilon_t \tag{5}$$

这里 W 是 $m \times n$ 矩阵。A 可以根据估计的 W 得到。如果 $m = n$, $A = W^{-1}$, 否则 A 是 W 的伪逆矩阵, $A = W'(WW')^{-1}$。

可以通过 PCA 得到白化数据 ε_t^w。令 V 表示白化矩阵, 其可以通过协方差矩阵 $E(\varepsilon_t\varepsilon_t')$ 的特征值分解得到: $V = ED^{-1/2}E'$，这里 D 是矩阵 $E(\varepsilon_t\varepsilon_t')$ 最大 m 个特征值组成的对角矩阵, 矩阵 E 的列是相应的特征向量。令 ε_t^w 表示白化误差, $\varepsilon_t^w = V\varepsilon_t$。很明显, $E(\varepsilon_t^w\varepsilon_t^{w'}) = I_m$, $W = \widetilde{W}V$。由于 $I_m = E(y_t y_t') = \widetilde{W}\widetilde{W}'$，所以 $\widetilde{W}$ 是正交矩阵。$\varepsilon_t^w = \widetilde{W}' y_t = \widetilde{W}'(H_t^y)^{1/2}\eta_t$。

Van der Weide (2002) 的 GO-GARCH 模型假设 h_{it}^y 服从经典的 $GARCH(1,1)$ 模型, 表示为:

$$h_{it}^y = (1 - \alpha_i - \beta_i) + \alpha_i y_{i,t-1}^2 + \beta_i h_{i,t-1}^y, \ \alpha_i, \beta_i \geqslant 0, \ \alpha_i + \beta_i < 1 \tag{6}$$

Fan 等(2008)[11] 建议采用更灵活的结构表示 h_{it}^y，认为 h_{it}^y 可能依赖于 $y_{j,t-k}, j \neq k, k \geqslant 1$。其扩展的 GARCH(1,1) 模型表示为:

$$h_{it}^y = (1 - \sum_{j=1}^m \alpha_{ij} - \beta_i) + \sum_{j=1}^m \alpha_{ij} y_{j,t-1}^2 + \beta_i h_{i,t-1}^y,$$

$$\alpha_{ij}, \beta_i \geqslant 0, \ \sum_{j=1}^m \alpha_{ij} + \beta_i < 1 \tag{7}$$

以上这些 GARCH 模型都忽略扰动项的符号, 因此是对称的方差过程。通常金融资产收益波动具有非对称和"杠杆效应"特征, 为此, 本文的扩展 GO-GARCH 模型分别假定因子 y_{it} 的方差 h_{it}^y 分别服从以下具有"杠杆效应"的 GJR、EGARCH 模型。

$GJR(P,Q)$ 模型的形式为:

$$h_{it}^y = K_i + \sum_{p=1}^P G_{ip} h_{i,t-p}^y + \sum_{q=1}^Q A_{iq} y_{i,t-q}^2 + \sum_{q=1}^Q L_{iq} S_{i,t-q} y_{i,t-q}^2 \tag{8}$$

如果 $y_{i,t-q} < 0, S_{i,t-q} = 1$, 表示"杠杆效应"。否则, $S_{i,t-q} = 0$。

为了满足方差 h_{it}^y 为正值和稳定性, 参数需要满足约束条件为:

$$\begin{aligned} &\sum_{p=1}^P G_{ip} + \sum_{q=1}^Q A_{iq} + \frac{1}{2}\sum_{q=1}^Q L_{iq} < 1 \\ &K_i > 0 \\ &G_{ip} > 0 \\ &A_{iq} > 0 \\ &A_{iq} + L_{iq} \geqslant 0 \end{aligned} \tag{9}$$

$GJR(P,Q)$ 中的 $\sum_{p=1}^P G_{ip} + \sum_{q=1}^Q A_{iq}$ 决定波动率持续性。

EGARCH(P, Q) 模型的形式为:

$$\log h_{it}^y = K_i + \sum_{p=1}^P G_{ip} \log h_{i,t-p}^y + \sum_{q=1}^Q A_{iq}\left[\frac{|y_{i,t-q}|}{\sqrt{h_{i,t-q}^y}} - E\left\{\frac{|y_{i,t-q}|}{\sqrt{h_{i,t-q}^y}}\right\}\right] + \sum_{q=1}^Q L_{iq}\left(\frac{y_{i,t-q}}{\sqrt{h_{i,t-q}^y}}\right) \tag{10}$$

对于正态分布, $E\{\eta_{t-q}\} = E\left\{\frac{|y_{i,t-j}|}{\sqrt{h_{i,t-q}^y}}\right\} = \sqrt{\frac{2}{\pi}}$。

EGARCH 和 GJR 模型是非对称方差模型, 能反映"杠杆效应", 或因子 y_{it} 与其波动率之间的负相关。如果存在"杠杆效应", 对于 GJR 模型, 杠杆系数 $L_{iq} > 0$，而对于 EGARCH 模型, 杠杆系数 $L_{iq} < 0$。可以证明, 当假定因子 y_{it} 方差分别服从具有以上非对称和"杠杆效应"的 GJR、EGARCH 模型时, 通过 $\varepsilon_t = Ay_t$ 线性变换, 条件协方差矩阵 H_t 也具有非对称和"杠杆效应", 这可以更好地反映现实金融资产收益波动性和相关性。

3 基于交互信息最小化的扩展广义正交 GARCH 模型参数估计

3.1 交互信息最小化与极大似然函数之间关系

交互信息来(mutual information)可以用来度量统计上的相关性。根据信息理论, m 个随机变量 $(y_1, \cdots, y_m)$ 之间的交互信息定义为:

$$I(y_1, \cdots, y_m) = \sum_{i=1}^m H(y_i) - H(y) \tag{11}$$

这里 $y = (y_1, \cdots, y_m)^T$，$H(\cdot)$ 表示(微分的)熵, 定义为: $H(x) = -\int p(x)\log p(x)\,dx$。通常 $I(y_1, \cdots, y_m)$ 是非负的, 当且仅当 y_i 彼此独立的时候为 0。与度量两个变量线性相关的线性相关系数

不同,交互信息不需要指定任何形式的相关,能同时捕捉到线性和非线性相关。

对于多元 GARCH 模型,令 η^* 是标准残差的集合:

$$\eta^* = (\eta_{1t}, \ldots, \eta_{mt}, \eta_{1,t-1}, \ldots, \eta_{m,t-1}, \ldots, \eta_{1,t-k}, \ldots, \eta_{m,t-k}) \tag{12}$$

这里 k 充分大。同样可以表示白化误差 ε^{w*} 的集合。

为了度量向量过程 $\eta \triangleq \{\eta_t\}$ 在同时期(contemporaneous dependence)和时序上(temporal dependence)的相关性,定义 η 的交互信息率为:

$$I(\eta) = \lim_{k\to\infty} \frac{1}{k+1} \cdot I(\eta^*) \tag{13}$$

可以通过使 $I(\eta)$ 最小化得到多元 GARCH 模型的参数估计。

根据 $\varepsilon^w_t = \widetilde{W}' y_t = \widetilde{W}'(H^y_t)^{1/2}\eta_t$。令 F 表示从 ε^{w*} 到 η^* 的映射,$\eta^* = F(\varepsilon^{w*}; \phi)$,这里 ϕ 表示参数集。可以计算 F 的 Jacobian 矩阵。

$$J = \begin{bmatrix} J_t \cdot \widetilde{W} & \cdots & \cdots & \cdots \\ 0 & J_{t-1} \cdot \widetilde{W} & & \\ \vdots & \vdots & & \\ 0 & \cdots & 0 & J_{t-k} \cdot \widetilde{W} \end{bmatrix} \tag{14}$$

这里:

$$J_t = \mathrm{diag}\{(h^y_{1t})^{-1/2}, (h^y_{2t})^{-1/2}, \ldots, (h^y_{mt})^{-1/2}\} \tag{15}$$

因为 $p_{\eta^*}(\eta^*) = \dfrac{p_{\varepsilon^{w*}}(\varepsilon^{w*})}{|J|}$, $|J| = \prod_{j=1}^{m}\prod_{i=0}^{k} h^{-1/2}_{j,t-i}$,所以有:

$$I(\eta) = \lim_{k\to\infty}\frac{I(\eta^*)}{k+1} = \lim_{k\to\infty}\frac{1}{k+1}\Big[\sum_{j=1}^{m}\sum_{i=0}^{k} H(\eta_{j,t-i}) - H(\eta^*)\Big]$$

$$= \lim_{k\to\infty} \frac{1}{k+1}\Big\{\sum_{i=0}^{k}\sum_{j=1}^{m} H(\eta_{j,t-i}) - [H(\varepsilon^{w*}) + E\log|J|]\Big\}$$

$$= \sum_{j=1}^{m}\Big[H(\eta_{jt}) + \frac{1}{2}E\log h^y_{jt}\Big] - \lim_{k\to\infty}\frac{1}{k+1}H(\varepsilon^{w*}) \tag{16}$$

以上方程最后一项与参数 ϕ 无关,因此,最小化 $I(\eta)$ 等价于最大化以下函数:

$$l(\phi) = \sum_{j=1}^{m}\Big[-H(\eta_{jt}) - \frac{1}{2}E\log h^y_{jt}\Big]$$

$$= \sum_{j=1}^{m} E\Big[\log p_{\eta_j}(\eta_{jt}) - \frac{1}{2}\log h^y_{jt}\Big] \tag{17}$$

因为假设 η_{jt} 是彼此独立,因此有:$p_\eta(\eta) = \prod_{j=1}^{m} p_{\eta_j}(\eta_{jt})$。由于 $\varepsilon^w = \widetilde{W}' y_t = \widetilde{W}'(H^y_t)^{1/2}\eta_t$,有 $p_\varepsilon(\varepsilon^w) = \dfrac{p_\eta(\eta_t)}{|\widetilde{W}'(H^y_t)^{1/2}|}$。这里 $H^y_t = \mathrm{diag}(h^y_{1t}, \ldots, h^y_{mt})$。因此,$\varepsilon^w_t$ 的对数似然函数为:

$$\log p_\varepsilon(\varepsilon^{w*}) = \sum_{t=1}^{T}\sum_{j=1}^{m}\Big[\log p_{\eta_j}(\eta_{jt}) - \frac{1}{2}\log h^y_{jt}\Big] \tag{18}$$

很明显,极大似然函数等价于最小交互信息。但采用极大似然估计 GO-GARCH 模型参数非常耗时,在数据维数很高时候会遇到一些困难。首先,在 GO-GARCH 估计过程中,不断更新的 W 和 GARCH 模型中的参数相互作用,可能导致估计困难。第二,极大似然函数值在最优点附近是非常平坦,不利于多元 GARCH 模型的估计。由于很难直接估计交互信息,可以采用一些统计准则近似最小化交互信息。作为选择,我们可以采用两个单独的步骤分别估计因子 y_t 和一元 GARCH 模型。首先,在本文提出的含"杠杆效应"的扩展 GO-GARCH 模型下通过最优化关于因子的一些统计准则确定 W 和因子 y_t。第二步,可以采用更灵活的 GJR 和 EGARCH 模型模拟因子 y_t。

3.2 基于不同统计技术实现交互信息最小化

(1)基于独立成分分析的因子提取方法

根据前面的分析可知,如果 GO-GARCH 被正确指定和估计,标准化残差 η_{jt} 在同时期和时序上一定是尽可能独立。由于 $\eta_t (t = k, k-1, \ldots,)$ 由 y_{jt} 决定,($\eta_t = y_{jt}(h^y_{jt})^{-1/2}$),与 η_i 无关($j \neq i$)。如果 η_{jt} 是时序独立的,因子 y_{jt} 一定是独立的。同时在弱假设条件下,可以通过使因子 y_{jt} 彼此尽可能独立而得到 W 和 y_{jt}。这与经典独立成分分析(ICA)方法目标完全一样(Hyvärinen 等 2001)[18]。

对于给定向量 ε,如果存在矩阵 A 和 m 维向量 $y_t = (y_{1t}, y_{2t}, \ldots, y_{mt})'$,其中对每个时刻 t, y_{it} 和 y_{jt} ($i \neq j$)都是相互独立的,使得 $\varepsilon = Ay_t$,则称 y_t 是 ε 的 m 个独立成分($m \leqslant N$)。A 称为转换矩阵或混合矩阵(mixing matrix)。考虑到可辨识性,通常假定各个独立成分满足:$E(y_t) = 0$, $E(y_t y_t') = I$。ICA 方法就是根据观测序列 ε 对这些潜在因子或变量 y_t,以及转换矩阵 A 进行估计。ICA 方法通过一定的计算方法得到矩阵 W,以便使得根据 $\hat{y}_t = W\varepsilon_t$ 估计的成分尽可能独立。然而,为了估计 A 和 y_t,需要满足至多一个独立成分是正态分布的假定

条件。由于对每个时刻 t, y_{it} 和 y_{jt} ($i \neq j$) 都是相互独立的, 收益率向量的条件方差 H_t 表示为: $H_t = AH_t^{y}A'$, 其中, $H_t^{y} = \mathrm{diag}\{h_{1t}^{y}, h_{2t}^{y}, \cdots, h_{mt}^{y}\}$, $y_{jt} = (h_{jt}^{y})^{1/2}\eta_{jt}$, $\eta_{jt} \sim i.i.d.(0,1)$, $j = 1, \cdots, m$。本文分别假定 h_{it}^{y} 服从 GARCH 模型、GJR 模型和 EGARCH 模型。最常用的估计独立成分方法限制 W 列为使 $\hat{y}_t$ 尽可能独立的方向。目前 ICA 计算的主要算法有 JADE (Cardoso 和 Souloumiac, 1993)[19]、FastICA (Hyvärinen, 1999)[20]、Infomax algorithm (Bell 和 Sejnowski, 1995)[21] 等。本文采用 FastICA 算法。

基于独立成分分析的扩展 GO-GARCH 模型, 与 O-GARCH 或 Van der Weide (2002)[10] 的 GO-GARCH 模型不同, H_t 的正定性能自动得到满足。而且, 模型参数估计可以分步完成, 并将协方差矩阵问题转化为多个一维问题来处理, 因此非常适合处理高维金融数据。基于独立成分分析的扩展 GO-GARCH 模型与 O-GARCH 或 GO-GARCH 模型最大的区别是 y_t 的成分相互独立, 从而保证其条件协方差矩阵 H_t 在构造上就是一个对角矩阵。这与 O-GARCH 或 Van der Weide (2002)[10] 的 GO-GARCH 模型中假设潜在因子或变量 y_t 的条件协方差矩阵是对角阵的处理方法不同。由于 ICA 的计算方法只能保证因子尽可能独立, 因子之间还不是完全独立。所以, 根据 ICA 方法估计 GO-GARCH 模型有可能得不到最优的协方差矩阵。另外, ICA 的计算方法仅仅利用因子之间的同时期的信息, 而没有考虑因子的时序信息。

(2) 基于因子条件不相关的因子提取方法

通常 GO-GARCH 模型假设因子 y_t 中成分是条件不相关的(conditionally uncorrelated), 以便 y_t 的条件协方差矩阵是对角阵。但是在计算中, 只是用因子间的无条件不相关(unconditional uncorrelatedness)代替条件不相关。其实, 因子间的无条件不相关不能确保其条件不相关, 这样可能导致在协方差矩阵估计中存在很大误差。为了减少误差, 一个自然的方法是执行因子间的条件去相关(conditional decorrelation)约束。Matsuoka 等(1995)[22] 研究表明, 如果因子是条件不相关的, 它们之间局部方差(local variances)变化是独立的, 因子(或矩阵 $\widetilde{W}$)能被唯一的确定。因此, 在施加条件去相关(conditional decorrelation)条件下, 可以对因子(或矩阵 W) 进行估计。Matsuoka 等 (1995)[22]、Hyvärinen 等(2001)[18] 在保证条件去相关条件下, 给出了估计因子 y_{jt} 的算法。但是这些算法需要在每次迭代之后, 重新计算因子的局部协方差, 这导致计算负担非常重。其实, 为了达到因子间的条件去相关, 在对数据白化之后, 正交矩阵 $\widetilde{W}$ 须使所有白化收益的局部协方差矩阵同时尽可能是对角阵。因此, 在估计一系列局部协方差矩阵之后, 可以通过同时对角化这些矩阵对 $\widetilde{W}$ 进行估计。本文采用 Cardoso 和 Souloumiac (1996)[23] 的同时对角化方法。局部协方差根据 EWMA 估计, 其平滑常数为 $\lambda = 0.94$。

为了表示方便, 本文分别把基于独立成分分析、基于因子条件不相关估计矩阵 $\widetilde{W}$ 和因子 y_t 方法简称为 ICA、CU。利用这些方法估计传统 GO-GARCH 模型简称为: ICA-GO-GARCH 模型、CU-GO-GARCH 模型; 用这些方法估计本文提出的含 "杠杆效应" 的扩展 GO-GARCH 模型简称为: ICA-GO-GJR 模型、CU-GO-GJR; ICA-GO-EGARCH 模型、CU-GO-EGARCH 模型。这些模型关注不同信息和利用不同目标函数, 因此, 得出不同因子 y_t 和矩阵 W。模型的选择依赖于哪种准则与金融数据特征最具有一致性。

4　实证研究

4.1　数据来源与处理

本文实证研究选择全球 15 个主要资本市场的股票指数作为研究对象。分别是: (1) 香港恒生指数 (HANG SENG Index)、(2) 上海证券交易所综合股价指数(SSE Composite Index)、(3) 深圳成分指数 (SSE COMPONENT Index)、(4) 日经指数 (NIKKEI)、(5) 新加坡海峡时报指数 (STRAITS TIMES Index)、(6) 韩国综合股价指数 (KOSPI Composite Index)、(7) 台湾证券交易所指数(TSEC weighted index)、(8) 巴黎 CAC40 指数、(9) 德国 DAX 综合指数、(10) 瑞士股票市场指数 (SSMI)、(11) 英金融时报指数(FTSE 100)、(12) 道琼斯指数 (Dow Jones Composite Average)、(13) 纽约股票指数(NYSE COMPOSITE INDEX)、(14) 纳斯达克指数(NASDAQ Composite)、(15) 标准普尔 500 指数 (S&P 500 INDEX)。时间从 2000 年 1 月 4 日—2009 年 11 月 13 日, 样本数据是调整后的收盘价格 (数据来源 Yahoo Finance)。由于不同国家的时差以及节假日不同, 我们对原始数据进行预处理, 最终每个市场得到 1997 个交易日的价格数据。然后将股票指数转换为日收益率 $r_t = \log(P_t / P_{t-1})$ 。各种

股票指数收益率的描述性统计和 GARCH 效应检验均可以明显看到所有收益序列都具有 GARCH 效应。另外,如果把各收益序列用图形表示也发现波动具有明显的"非对称"和"杠杆效应"特征。

4.2 实证估计结果与分析

根据样本数据,对基于交互信息最小化估计含"杠杆效应"的扩展 GO-GARCH 模型实证研究。实证中同时选择对不同类型的多元 GARCH 模型进行比较与分析。这些模型包括: O-GARCH (Alexander, 1997)[9] 模型、DCC 模型(Engle, 2002)[5]、GO-GARCH 模型 (Van der Weide, 2002)[10]。没有选择其他模型的原因是: 其他模型如 VECH 模型 (Bollerslev 等, 1988)[1] 和 BEKK 模型(Engle 和 Kroner, 1995)[2] 中包含太多的参数,不适合模拟高维数据的条件协方差矩阵。由于篇幅限制,省略各个模型参数和条件协方差矩阵 H_t,需要者可以向作者索取。根据不同模型得到的条件相关性差别很大。可以通过假设检验来分析不同类型多元 GARCH 模型对协方差矩阵 H_t 的拟合程度。

根据 Tse (1999)[24]、Engle(2002)[5],可以采用 Box-Pierce 基于残差交叉乘积的检验方法。具体来说,设第 i 序列的标准化残差为: $\eta_{it} = \varepsilon_{it}/h_{ii,t}^{1/2}$, $i = 1, \cdots, N$。定义如下统计量:

$$C_{ij,t} = \begin{cases} \eta_{it}^2 - 1, & i = j; \\ \eta_{it}\eta_{jt} - \rho_{ij,t}, & i \neq j \end{cases} \tag{19}$$

其中: $\rho_{ij} = h_{ij,t}/\sqrt{h_{ii,t}h_{jj,t}}$ 表示 ε_{it} 和 ε_{jt} 之间的相关系数。如果 H_t 的模型选择合适,那么对于任何 i 和 j, $C_{ij,t}$ 都应没有任何自相关性。可以用 Box-Pierce 的 Q 统计量检验这种相关性:

$$Q(i,j;M) = T\sum_{r=1}^{M} R_{ij,r}^2 \tag{20}$$

$R_{ij,t}$ 是 $C_{ij,t}$ 的第 r 个样本自相关函数。可以采用 $\chi^2(M)$ 分布检验自相关函数都为零的假设[25]。

本文同时也应用 Box-Pierce/Ljung-Box 的 portmanteau 检验模型的拟合程度。根据 Hosking (1980)[26],可以给出多元形式的 Ljung-Box 检验统计量:

$$HM(M) = n^2\sum_{j=1}^{M}(n-j)^{-1}tr\{C_{Z_t}^{-1}(0)C_{Z_t}(j)C_{Z_t}^{-1}(0)C'_{Z_t}(j)\} \tag{21}$$

这里: $Z_t = vech(\eta_t\eta'_t)$, $\eta_t = H_t^{-1/2}\varepsilon_t$。$C_{Z_t}(j)$ 是样本 j 阶自协方差矩阵, $C_{Z_t}(j) = n^{-1}\sum_{t=j+1}^{n}(Z_t - \overline{Z})(Z_t - \overline{X})'$, $j = 0, \cdots, n-1$。$\overline{Z} = (Z_1 + \cdots + Z_n)/n$。在原假设为不存在 ARCH 效应条件下, $HM(M)$ 的渐近分布为 $\chi^2(K^2M)$,常数 M 是决定检验统计的滞后量。表 1 为根据不同模型得到的 $Q(i,j;M)$,表中的*、**、*** 分别表示 0.1, 0.05, 0.001 显著性水平;这里 $1 \leqslant i \leqslant j \leqslant 15$, $M = 5$。显著性水平是根据 χ_5^2 分布计算得到。为了节省篇幅,这里只是给出部分结果,即表示模型对主要指数收益波动率和相关性反映。这些指数是: (1) 香港恒生指数(HANG SENG Index); (2) 上海证券交易所综合股价指数(SSE Composite Index); (4) 日经指数(NIKKEI); (9) 德国 DAX 综合指数; (11) 英金融时报指数 (FTSE 100); (14) 纳斯达克指数 (NASDAQ Composite); (15) 标准普尔 500 指数 (S&P 500 INDEX)。表 2 给出 $1 \leqslant M \leqslant 10$ 时的 $HM(M)$ 值。$HM(M)$ 值越大表示相关的模型拟合效果越差。

表 1 不同模型得到的 **$Q(i, j, M)$**

i, j	DCC	O-GARCH	GO-GARCH	ICA-GO-GARCH	ICA-GO-GJR	ICA-GO-EGARCH	CU-GO-GARCH	CU-GO-GJR	CU-GO-EGARCH
1, 1	2.32	30.10***	4.81	23.64***	32.76***	34.29***	7.99	6.65	2.12
1, 2	1.37	0.75	0.84	0.53	0.84	1.97	0.68	0.85	0.72
1, 4	14.54**	8.75	8.33	9.69	8.96	7.03	8.63	7.51	5.54
1, 9	2.22	8.62	4.68	10.01	7.78	18.11***	6.73	6.52	4.82
1, 11	3.23	1.05	2.07	2.43	1.52	1.16	3.26	2.88	3.4
1, 14	11.51**	8.76	5.13	5.13	6.89	6.35	6.64	6.83	7.59
1, 15	16.14***	5.53	6.79	4.46	1.84	7.74	5.05	6.93	4.12
2, 2	1.43	11.59***	3.71	4.87	6.5	4.75	4.12	4.58	0.97
2, 4	1.54	7.32	3.59	4.78	4.16	5.73	5.04	4.88	4.8
2, 9	0.23	2.13	1.05	3.35	2.57	1.7	1.62	1.54	1.72
2, 11	2.87	34.96***	14.98***	15.66***	14.75**	13.36**	9.38	9.57	8.5
2, 14	3.63	2.25	0.51	0.53	0.26	3.27	0.6	0.5	1.29
2, 15	1.97	17.83***	5.34	8.11	8.43	6	5.85	5.85	4.71

续

i, j	DCC	O-GARCH	GO-GARCH	ICA-GO-GARCH	ICA-GO-GJR	ICA-GO-EGARCH	CU-GO-GARCH	CU-GO-GJR	CU-GO-EGARCH
4, 4	5.48	7.59	3.94	4.03	4.45	6.18	3.82	3.82	3.44
4, 9	4.4	0.75	1.95	0.88	1.56	0.65	1.32	1.52	0.82
4, 11	3.35	6.28	4.02	4.88	5.55	3.39	2.8	3.45	2.96
4, 14	0.1	14.75^{**}	3.9	8.35	8.79	12.39^{**}	5.51	5.65	7.6
4, 15	2.36	8.41	3.93	2.47	3.85	4.19	5.87	5.3	5.91
9, 9	2.26	35.26^{***}	0.96	16.11^{**}	5	1.21	1.01	1.22	0.85
9, 11	4.46	7.35	3.45	5.97	5.67	3.3	4.21	4.78	3.82
9, 14	3.14	21.10^{***}	10.81^{*}	7.91	5.69	6.42	7.64	7.26	7.16
9, 15	2.35	23.89^{***}	2.19	2.04	2.3	1.22	1.28	1.78	0.82
11, 14	2.51	9.8	2.86	2.44	1.84	5.03	4.1	4	4.63
11, 15	11.76^{**}	5.19	5.61	11.64^{**}	7.86	11.40^{**}	8.1	6.79	8.39
14, 14	0.61	167.42^{***}	3.94	7.3	7.74	7.79	5.15	5.73	7.03
14, 15	4.78	36.41^{***}	6.59	8.33	8.06	6.18	5.14	3.56	6.12

表 **2**　不同模型的 **HM(M)** 统计

k	DCC	O-GARCH	GO-GARCH	ICA-GO-GARCH	ICA-GO-GJR	ICA-GO-EGARCH	CU-GO-GARCH	CU-GO-GJR	CU-GO-EGARCH
1	19917	19776	17913	18622	18367	18771	18256	18217	17895
2	40617	39572	36633	37718	37188	38397	37208	36996	36856
3	55761	55805	52959	54120	53528	54731	53677	53558	53206
4	72810	73269	69847	71595	71109	72061	70764	70281	69932
5	90454	91062	86845	88831	88603	89621	87804	87334	87250
6	105779	106295	102069	104402	104271	105649	103038	102433	102529
7	122584	123446	118557	121137	120830	123068	119384	118787	119737
8	138451	140642	135047	137853	137596	140383	136029	135458	136777
9	154761	157161	151118	154233	154013	157486	152168	151584	153465
10	171180	174248	167464	171314	171220	175726	168939	168359	171312

根据计算结果, 可以发现:

(1) $Q(i,j;M)$ 值和 $HM(M)$ 值表明, DCC 模型、O-GARCH 模型的效果最差。对于 DCC 模型, $Q(1,4;5)=14.54$、$Q(1,14;5)=14.54$, $Q(1,15;5)=16.14$, $Q(1,15;5)=11.76$, 分别在 0.05、0.001、0.01 的显著水平下拒绝模型。而对于 O-GARCH 模型, $Q(1,1;5)=30.10$、$Q(2,11;5)=34.96$、$Q(2,15;5)=17.83$、$Q(4,14;5)=14.75$、$Q(9,9;5)=35.26$、$Q(9,14;5)=21.10$、$Q(9,15;5)=23.89$、$Q(14,14;5)=167.42$、$Q(14,15;5)=11.59$, 至少在 0.001 显著水平下拒绝模型。

(2) $Q(i,j;M)$ 值和 $HM(M)$ 值表明, Van der Weide(2002)[10] 的传统基于极大似然估计的 GO-GARCH 模型要比 DCC 模型、O-GARCH 模型的效果要明显好。对于 Van der Weide(2002)[10] 的基于极大似然估计的 GO-GARCH 模型, 只有 $Q(2,11;5)=14.98$, $Q(9,14;5)=10.81$, 分别在 0.1、0.05 显著水平下拒绝模型。但 Van der Weide (2002)[10] 基于极大似然估计的 GO-GARCH 模型最大的缺陷是计算速度非常慢, 对于本文 15 维的运算, 参数估计大约需要 8 个小时。

(3) $Q(i,j;M)$ 值和 $HM(M)$ 值表明, ICA-GO-GARCH 模型绩效并不比传统基于极大似然估计的 GO-GARCH 模型绩效高。分别有多个 $Q(i,j;5)$ 值至少在 0.05, 0.01 水平下显著拒绝模型。

(4) 根据 $Q(i,j;M)$ 值, CU-GO-GARCH 模型、CU-GO-GJR 模型、CU-GO-EGARCH 模型绩效比传统基于极大似然估计的 GO-GARCH 模型绩效高, 都通过统计检验。同时根据 CU-GO-GJR 模型、CU-GO-EGARCH 模型得到的 $HM(M)$ 值较低, 这不仅表明采用基于因子条件不相关方法估计 GO-GARCH 模型比较合适, 而且说明不能忽视因子波动存在的"非对称"和"杠杆效应"。

5　结语

本文对 Van der Weide (2002)[10] 最初的 GO-GARCH 模型的进行扩展, 建立反映金融资产收益的波动特征, 具有"非对称"和"杠杆效应"特征的扩展广义正交 GARCH 模型。并从交互信息理论视

角,在理论上分析了基于交互信息最小化与基于极大似然函数参数估计的多元 GARCH 模型参数估计方法联系和区别。这种基于交互信息最小化的参数估计方法既保持了极大似然参数估计的优点,同时也避免传统极大似然函数估计需要事先正确指定标准化残差概率密度函数的困难。本文同时在扩展广义正交 GARCH 模型框架下,采用不同的统计技术实现基于交互信息最小化的参数估计方法,这些方法利用数据的不同信息来估计转换矩阵和提取因子,模型参数能被快速可行估计,计算效率非常高,解决了多元 GARCH 模型的"高维数灾难"问题,使其本文提出的扩展广义正交 GARCH 模型可以在高维系统中应用。

本文分别对不同多元 GARCH 模型和基于交互信息最小化的扩展广义正交 GARCH 模型进行实证研究。结果显示,DCC 模型、O-GARCH 模型的效果相对较差。Van der Weide (2002)[10] 基于极大似然估计的 GO-GARCH 模型比 DCC 模型、O-GARCH 模型的效果要明显好。但 Van der Weide (2002)[10] 基于极大似然估计的 GO-GARCH 模型最大的缺陷是计算速度非常慢,这种"高维数灾难"直接限制了其在实际中的应用。研究发现基于传统独立成分分析(ICA)的各种 GO-GARCH 模型尽管计算速度较快,但是模型绩效并不比传统基于极大似然估计的 GO-GARCH 模型绩效高。而基于因子条件不相关方法(CU)估计,假设因子服从 EGARCH、GJR 的扩展广义正交 GARCH 模型绩效较高,而且计算速度较快。这不仅表明采用基于因子条件不相关方法(CU)估计 GO-GARCH 模型比较合适,而且说明不能忽视因子波动存在的"非对称"和"杠杆效应"。可以采用这些方法对资产定价、资产组合选择、期权定价、套期保值、风险管理等问题进行进一步研究。

本文统计检验方法是针对历史数据得到的结论,表明模型对各种股票指数收益的波动性和协同波动性的拟合效果。将来还需要对这些模型进行事后样本(post-sample)比较,对模型的预测能力进行评价。由于条件协方差矩阵属于隐含变量,需要构建条件协方差矩阵的代理变量(proxies)。另外,本文只是假设线性映射矩阵不发生变化为前提,而对于时变线性映射矩阵研究没有过多探讨,这非常值得在以后的研究中加强。

参考文献:

[1] Bollerslev, T., Engle, R. F., Wooldridge, J. M.. A capital asset pricing model with time-varying covariances [J]. The Journal of Political Economy, 1988, 96: 116-131.

[2] Engle, R. F., Kroner, K. F.. Multivariate simultaneous generalized ARCH [J]. Econometric Theory, 1995, 11: 122-150

[3] Bollerslev, T.. Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model [J]. Review of Economics and Statistics, 1990, 72: 498-505.

[4] Tse, Y. K., Tsui, K. C.. A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations [J]. Journal of Business and Economic Statistics, 1999, 20: 351-362

[5] Engle, R.. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models [J]. Journal of Business and Economic Statistics, 2002, 20: 339-350.

[6] Engle, R. F., Ng, V. K., Rothschild, M.. Asset pricing with a factor ARCH covariance structure: Empirical estimates for treasury bills [J]. Journal of Econometrics, 1990, 45: 213-238

[7] 刘志东. 多元 GARCH 模型结构特征、参数估计与假设检验研究[J]. 数量经济技术经济研究, 2010, 27(9): 147-161.

[8] Ding, Z.. Time Series Analysis of Speculative Returns [D]. PhD Dissertation, University of California at San Diego, 1994

[9] Alexander, C., Chibumba, A.. Multivariate Orthogonal Factor GARCH [Z]. University of Sussex, Mimeo, 1997.

[10] Van der Weide, R.. GO-GARCH: A multivariate generalized orthogonal GARCH model [J]. Journal of Applied Econometrics, 2002, 17: 549-564.

[11] Fan, J., Wang, M., Yao, Q.. Modeling multivariate volatilities via conditionally uncorrelated components [J]. Journal of the Royal Statistical Society Series B, 2008, 70: 679-702

[12] Boswijk, H. P., Van der Weide, R.. Wake me up before you GO-GARCH [Z]. Tinbergen Institute Discussion Paper, TI 2006-079/4, University of Amsterdam and Tinbergen Institute.

[13] Broda, S. A., Paolella, M. S.. CHICAGO: A fast and accurate method for portfolio risk calculation [J]. Journal of Financial Econometrics, 2009, 7(4): 412-436.

[14] 王明进、陈奇志. 基于独立成分分解的多元波动率预测[J]. 管理科学学报, 2006, 9(5): 56-64.

[15] 王明进. 高维波动率预测[J]. 数量经济技术经济研究, 2008, 11: 137-148.

[16] 许启发、张世英. 多元条件高阶矩波动性建模[J]. 系统工程学报, 2007, 22(1): 1-9.

[17] Zhang, K., Chan, L. W.. Efficient factor GARCH models and factor-DCC models [J]. Quantitative Finance, 2009, 9(1): 71-91.

[18] Hyvärinen, A., Karhunen, J., Oja, E.. Independent Component Analysis [M]. UK: John Wiley & Sons, Inc, 2001.

[19] Cardoso, J. F., Souloumiac, A.. Blind beamforming for non-Gaussian signals [C]. IEE Proceedings F, 1993, 140: 362-370

[20] Hyvärinen, A., Oja, E.. A fast fixed-point algorithm for independent component analysis [J]. Neural Computation, 1997, 9: 1483-1492

[21] Bell, A. J., Sejnowski, T. J.. An information-maximization approach to blind separation and blind deconvolution [J]. Neural Computation, 1995, 7: 1129-1159.

[22] Matsuoka, K., Ohya, M., Kawamoto, M.. A neural net for blind separation of nonstationary signals [J]. Neural Networks, 1995, 8: 411-419.

[23] Cardoso, J. F., Souloumiac, A.. Jacobi angles for simultaneous diagonalization [J]. SIAM Journal on Matrix Analysis and Applications, 1996, 17: 161-164

[24] Tse, Y., Tsui, A.. A note on diagnosing multivariate conditional heteroscedasticity models [J]. Journal of Time Series Anaysis, 1999, 20: 679-691.

[25] Tse, Y.. Residual-based diagnostics for conditional heteroscedasticity models [J]. Econometrics Journal, 2002, 5: 358-373

[26] Hosking, J.. The multivariate portmanteau statistic [J]. Journal of American Statistical Association, 1980, 75: 602-608

An Extended Generalized Orthogonal GARCH Models for Multivariate Volatility in High Dimensional System of Finance Market and Its Parameter Estimation

LIU Zhi-dong, XUE Li

(1. Central University of Finance and Economics, Beijing 100081, China;
2. Nan Jing University, Nanjing 210093, China)

Abstract: In this paper, an extended generalized orthogonal GARCH Model, which can reflect the asymmetry and leverage of volatility, is proposed based on that of Van der Weide (2002). Then, a new estimation method for the multivariate GARCH models from the mutual information-theoretic viewpoint is considered. The relationship between the statistical dependence in standardized residuals and the maximum likelihood, when estimating multivariate GARCH models, is revealed. Based on that, the different statistic ways are proposed in the framework of the extended generalized orthogonal GARCH models, for purpose of implementing the models based on the estimation of mutual information minimization. These methodologies can therefore be easily applied to high-dimensional systems, where likelihood-based estimation will run into computational problems. According to 15 global stock returns indexes, empirical research is included to illustrate the model and the estimation method.

Key words: mutual information; multivariate GARCH models; leverage effects; parameter estimation; dynamic conditional correlation