中国科学 A 辑: 数学 2009 年 第 39 卷 第 7 期: 801～816
www.scichina.com math.scichina.com
《中国科学》杂志社 SCIENCE IN CHINA PRESS

时间序列中向量自回归模型的拟合优度检验

吴鑑洪 ①*, 朱力行 ②

① 浙江工商大学统计与数学学院, 杭州 310018
② 香港浸会大学数学系, 香港
E-mail: wjhstat1@yahoo.com.cn

收稿日期: 2007-07-29; 接受日期: 2009-02-02; * 通信作者
香港研究资助局 (批准号: HKBU2030/07P) 和中国教育部人文社科 (批准号: 07JJD790154) 以及浙江工商大学青年人才基金 (批准号: Q09-12) 资助项目

摘要 本文提出并研究了一些诊断检验工具, 用于一般参数型的时间序列向量自回归模型的拟合优度检验. 该检验在零假设下渐近服从卡方分布, 并能侦察到以参数速度收敛到零假设模型的备择模型. 检验涉及到权函数, 因此可以灵活地选择权函数以提高检验功效, 尤其是在可能的偏离方向已知情形. 如果备择不是方向型的, 而只知道其属于某一个模型类中, 此时可构造一个渐近分布自由的极大极小 (maximin) 检验. 对于饱和备择, 基于得分型思想给出了构造万能 (omnibus) 检验的可行性构想. 本文对提出的检验从理论上进行了功效研究. 另外, 为提高检验在小样本情形的功效, 本文把非参数 Monte Carlo 检验方法推广到相依数据情形. 最后, 通过模拟研究和实际数据分析进一步表明检验的有用性.

关键词 拟合优度检验 maximin 检验 非参数 Monte Carlo 检验 得分型检验 时间序列 向量自回归模型

MSC(2000) 主题分类 62F05, 62H15

1 引言

近年来, 信息和通讯技术的飞速发展以及经济全球化加快了世界金融市场一体化的进程. 随之而来的是金融市场间的相互依赖关系比以前更强了. 为了更好的理解全球金融的动态结构, 研究分析多个金融市场间的相互关系是必要的和重要的. 比如, 对一个拥有多个资产的投资者或金融机构来说, 掌握金融资产间的收益率的动态关系对制定决策是至关重要的. 众所周知, 多个资产的收益率可通过多元 (或向量) 时间序列统计建模和统计推断来进行研究分析, 更多细节可参考专著 [1]. 尽管多元时间序列在实际应用中很重要, 但是文献中相关的研究不多见. 文献中许多工作仍然是关于单个时间序列的统计分析和推断.

本文考虑一般的向量自回归 (GVAR) 模型

$$X_t = \mu(F_{t-1}) + \varepsilon_t,$$

引用格式: 吴鑑洪, 朱力行. 时间序列中向量自回归模型的拟合优度检验. 中国科学 A, 2009, 39(7): 801–816
Wu J H, Zhu L H. Goodness-of-fit tests for vector autoregressive models in time series. Sci China Ser A, 2009, 52, DOI: 10.1007/s11425-009-0071-1

其中 X_t 是 k 元时间序列, F_{t-1} 是由 $\{X_{t-1}, X_{t-2}, \ldots\}$ 产生的 σ 域, 上式右边第一项是 X_t 关于 σ 域 F_{t-1} 的自回归函数, 误差项 ε_t 正交于 F_{t-1}, 即 $E(\varepsilon_t|F_{t-1}) = 0$. 由于在实际应用中自回归函数 $\mu(\cdot)$ 是未知的, 关于回归函数 $\mu(\cdot)$ 的统计推断是一个很重要的问题, 并得到人们越来越多的关注. 一个常用的方法是假设隐含的回归模型属于一个参数族 $\mathcal{M}=\{m(\theta,\cdot):\theta\in\Theta\subseteq\mathbb{R}^p\}$, 如此假设的原因之一是参数族模型能够以简洁明了的方式表现过去及现在的观测数据对未来预测变量的影响, 并且处理起来比较简单.

在著名的 Box-Jenkins 时间序列建模方法论中, 许多工作致力于模型的提出以及参数估计, 模型诊断检验没有得到相应的重视, 尽管三者从理论上来讲是同等的重要. 本文提出并研究了一些诊断检验工具用于参数型向量自回归模型的拟合优度检验. 具体而言, 我们检验 $\mu(\cdot)$ 是否在函数类 $\mathcal{M}$ 中, 其中 $\mathcal{M}=\{m(\theta,\cdot):\theta\in\Theta\subseteq\mathbb{R}^p\}$ 是给定的某个参数函数类. 诚然, 一元时间序列模型的模型检验方法经过适当的修改可以用来处理多元情形的相应问题. 但是, 在研究多元时间序列数据时, 要特别地关注因变量各成分间的相关性. 另外, 已知方法的直接推广常常得不到功效高的检验.

本文研究了形式一般的多元时间序列自回归模型的检验问题, 通过对残差各分量逐一加权平均得到得分型拟合优度检验. 检验统计量的极限分布可用于确定临界值. 该检验对模型的错误指定是敏感的, 能侦察到以最快的速度 $(n^{-\frac{1}{2}})$ 收敛到零模型的备择模型, 并可通过选择适当的权函数构造功效高的检验. 然而当备择不是方向型时, 单个的权函数常常不管用, 因为单个权函数一旦选择不当会大大降低检验的功效. 更常见的情形是人们可能知道某个函数类包含了隐含的备择. 此时, 在构造检验时就应该考虑这些可能的偏离, 充分利用好已知信息. 更多细节, 读者可参考文献 [2–4]. 本文针对这种包含若干个可能的备择模型的备择类并基于得分型检验思想构造了渐近分布自由的极大极小 (maximin) 检验, 相应的得分与可能的偏离有关以便构造出的得分型检验具有高的功效.

当没有任何关于备择模型的相关信息时, 也即所谓的饱和备择情形, 检验问题则是完全非参数的, 需要构造一个万能 (omnibus) 检验来处理这个问题. 著名的混成 (portmanteau) 检验不是万能检验, 不能处理这个问题, 见文献 [5]. 注意到本文提出的检验由于引入得分构造检验, 通过非正式的残差图帮助我们选择权函数提高检验的功效. 这是该检验的一大优点, 尽管某种意义上增加了权函数的选择问题. 并进一步讨论了如何基于得分思想构造极大极小检验部分地处理饱和备择问题. 本文第 5 节也特别讨论了基于得分思想构造万能检验的可行性方案. 不过, 这并不意味我们彻底解决了这个问题. 这个问题值得进一步研究.

文献中有相当一部分工作是构造万能 (omnibus) 检验来处理备择模型完全未知之情形, 比如文献 [6–12]. 上述方法都是针对一元因变量情形, 而且回归变量也是一元的. 另外, 尽管万能检验可以用来检验饱和的备择, 但是当备择模型的一些先验信息可用时检验的功效相对而言蒙受损失. 到目前为止, 文献中还没有人提出万能检验来对多元时间序列自回归模型进行模型检验. 作者认为当备择的一些先验信息已知时, 应该利用好这些已知信息构造功效高的检验. 本文构造的得分型检验充分利用了先验信息, 具有较高的功效.

另外, 在古典统计中, 非参数 Monte Carlo 检验 (NMCT) 方法得到很好的研究. 研究表明当样本量比较小时, 该检验方法非常有用, 详情可参考专著 [13]. 本文把这种方法推广到相依数据情形, 并比较了用检验的极限分布和用非参数 Monte Carlo 检验 (NMCT) 方法确定临界值的检验功效. 结果表明当样本量大或者适中时, 用极限分布确定临界值表现比较好,

此时非参数 Monte Carlo 检验 (NMCT) 方法并未有更多优势; 但是当样本量很小时, 用极限分布确定临界值检验功效比较低, 而用非参数 Monte Carlo 检验 (NMCT) 方法明显比用极限分布的方法好.

本节最后给出后面要用到的一些记号. 对矩阵 A, A^{T} 表示其转置, A^{-1} 表示其逆矩阵, $\mathrm{vec}(A)$ 表示其通常意义下的拉直. 符号 "$\xrightarrow{D}$" 表示依分布收敛, "$\sum$" 表示沿着指标 t 的求和, 即 "$\sum_t$". $a_n = o_p(b_n)$ 表示 a_n/b_n 依概率收敛到零, $a_n = O_p(b_n)$ 表示 a_n/b_n 依概率有界. 对 $k\times 1$ 向量 a, b 和 $k\times q$ 矩阵 A, 定义点乘运算为 $a\bullet b = \mathrm{diag}(a)b$, $a\bullet A = A\bullet a = \mathrm{diag}(a)A$, 其中 $\mathrm{diag}(a)$ 表示以 a 中第 i 元素作为第 i 个对角元素的 $k\times k$ 对角矩阵. $A\otimes B$ 表示矩阵 A 和 B 的 kronecker 积.

2 检验的构造和功效研究

2.1 检验的构造

设 $X_t = (X_{1t},\ldots,X_{kt})^{\mathrm{T}}$ 是平稳遍历的 k 维时间序列, 考虑如下参数向量自回归模型

$$H_0:\ X_t = m_t + \varepsilon_t, \tag{1}$$

其中 $m_t = m(\theta, F_{t-1})$, ε_t 是条件均值为零, 条件方差为 $V_t = V(\theta, F_{t-1})$ 的误差项, F_t 是由 $\{X_t, X_{t-1},\ldots\}$ 产生的 σ 域. 不妨设 $\varepsilon_t = V_t^{\frac{1}{2}}\eta_t$, 标准化的误差 $\eta_t = (\eta_{1t},\ldots,\eta_{kt})^{\mathrm{T}}$ 是均值为零、方差阵为 I_k (单位阵) 的独立同分布序列, 并独立于 $\{X_{t-1}, X_{t-2},\ldots\}$. 正是 η_t 的这些性质启发我们构造合适检验统计量检验模型 (1) 的拟合优度. 假设不计未知的参数, 函数 m_t 和 V_t 是已知的, 且有连续的二阶偏导.

在零假设模型 (1) 下, 对任何 k 维 F_{t-1} 可测的权函数 $\omega(\theta, F_{t-1})$ (简记为 ω_t), 若期望存在, 则有 $E(\eta_t\bullet\omega_t) = 0$. 并根据其经验形式定义一个得分型检验

$$T_{n0} = \frac{1}{\sqrt{n}}\sum \hat{\eta}_t\bullet\hat{\omega}_t,$$

其中 $\hat{\eta}_t = \hat{V}_t^{-\frac{1}{2}}(X_t - \hat{m}_t)$ 是相应的残差, $\hat{V}_t$ 和 $\hat{m}_t$ 分别是 V_t 和 m_t 的估计, $\hat{\omega}_t$ 类似定义. 为了得到标准化的检验, 定义二次型检验统计量 $T_n^2 = T_{n0}^{\mathrm{T}}\Sigma_n^{-1}T_{n0}$, 其中 Σ_n 是 T_{n0} 的渐近协方差估计.

作为检验步骤的一部分, 未知参数的估计是一个重要的问题. 当新禧过程 ε_t 是条件正态时, 可通过使其对数似然函数极大化

$$l = \sum l_t, \qquad l_t = -\frac{1}{2}\log|V_t| - \frac{1}{2}\varepsilon_t^{\mathrm{T}}V_t^{-1}\varepsilon_t, \tag{2}$$

得到所谓的极大似然估计 $\hat{\theta}$. 然而在许多实际应用中, 很难保证满足新禧的条件正态性假设. 本文不假设这个条件正态性, 但仍然通过使上面似然函数 l 极大化来求参数的估计. 文献中称这样得到的估计为拟极大似然估计 (QMLE). 在正确指定的自回归模型 (1), 以及一些正则条件 (比如文献 [14, p. 167] 给出的条件 A.1) 下,

$$n^{\frac{1}{2}}(\hat{\theta}-\theta) = \frac{1}{\sqrt{n}}J^{-1}\sum\frac{\partial l_t}{\partial\theta} + o_p(1), \tag{3}$$

$$n^{\frac{1}{2}}(\hat{\theta}-\theta) \xrightarrow{D} N(0, J^{-1}IJ^{-1}), \tag{4}$$

其中 $I = E(\frac{1}{n}\sum\frac{\partial l_t}{\partial\theta}\frac{\partial l_t}{\partial\theta^{\mathrm{T}}})$, $J = E(-\frac{1}{n}\sum\frac{\partial^2 l_t}{\partial\theta\partial\theta^{\mathrm{T}}})$.

在零假设模型 (1) 和第 6 节条件 A.1–A.3 下, 可以证明 (细节见第 6 节),

$$T_{n0}=\frac{1}{\sqrt{n}}\sum\eta_t\bullet\omega_t-\frac{1}{\sqrt{n}}A^{\mathrm{T}}J^{-1}\sum\frac{\partial l_t}{\partial\theta}+o_p(1), \tag{5}$$

其中 $A=E[\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-\frac{1}{2}}\mathrm{diag}(\omega_t)]$. 注意到上式中 T_{n0} 的主要部分是鞅差序的和的形式. 在平稳性和遍历性假定下, 根据鞅中心极限定理 (见文献 [15, p. 788]) 和 Cramér-Wold 技术, 可得

$$T_{n0}\xrightarrow{D}N(0,\Sigma),$$

其中

$$\begin{aligned}\Sigma&=E\left[\left(\eta_t\bullet\omega_t-A^{\mathrm{T}}J^{-1}\frac{\partial l_t}{\partial\theta}\right)\left(\eta_t\bullet\omega_t-A^{\mathrm{T}}J^{-1}\frac{\partial l_t}{\partial\theta}\right)^{\mathrm{T}}\right]\\&=\Sigma_0+A^{\mathrm{T}}J^{-1}IJ^{-1}A-2A^{\mathrm{T}}J^{-1}A,\\\Sigma_0&=E[\mathrm{diag}(\omega_t)\mathrm{diag}(\omega_t)].\end{aligned}$$

于是检验 $T_n^2=T_{n0}^{\mathrm{T}}\Sigma_n^{-1}T_{n0}$ 渐近服从卡方分布, 其中 Σ_n 是 T_{n0} 渐近方差 Σ 的一个相合估计,

$$\Sigma_n=\frac{1}{n}\sum[\mathrm{diag}(\hat\omega_t)\mathrm{diag}(\hat\omega_t)]+\hat A^{\mathrm{T}}\hat J^{-1}\hat I\hat J^{-1}\hat A-2\hat A^{\mathrm{T}}\hat J^{-1}\hat A, \tag{6}$$

$$\hat A=\frac{1}{n}\sum\left(\frac{\partial\hat m_t^{\mathrm{T}}}{\partial\theta}\hat V_t^{-\frac{1}{2}}\mathrm{diag}(\hat\omega_t)\right),\ \hat I=\frac{1}{n}\sum\frac{\partial\hat l_t}{\partial\theta}\frac{\partial\hat l_t}{\partial\theta^{\mathrm{T}}},\ \text{和}\ \hat J=-\frac{1}{n}\sum\frac{\partial^2\hat l_t}{\partial\theta\partial\theta^{\mathrm{T}}}.$$

定理 1 在零假设模型 (1) 和第 6 节条件 A.1–A.3 下,

$$T_n^2\xrightarrow{D}\chi_k^2,$$

其中 χ_k^2 是自由度为 k 的卡方分布.

2.2 功效研究和更多检验的构造

本小节研究检验 T_n^2 在备择模型下的渐近性质. 考虑一序列与 n 有关的模型

$$H_{1n}:X_{nt}=m_t+C_ns_t+\varepsilon_t, \tag{7}$$

$t=1,2,\ldots,n$, $F_{n,t-1}$ 可测函数 $s_t=(s_{1t},\ldots,s_{kt})^{\mathrm{T}}$ 不在参数类 $\mathcal{M}=\{m(\theta,\cdot):\theta\in\Theta\subseteq\mathbb{R}^p\}$ 中, C_n 是常数序列, $m_t=m(\theta,F_{n,t-1})$ 和 ε_t 与上文一致. 我们可得到如下定理.

定理 2 假设第 6 节条件 A.1–A.5 满足. 在备择假设模型 (7) 下, 若 $n^\gamma C_n\to r$ $(r\neq 0), 0\leqslant\gamma<\frac{1}{2}$, 则依概率

$$T_n^2\longrightarrow\infty,$$

若 $n^{\frac{1}{2}}C_n\to 1$, 则

$$T_n^2\xrightarrow{D}\chi_k^2(C^{\mathrm{T}}\Sigma^{-1}C),$$

其中 $\chi_k^2(C^{\mathrm{T}}\Sigma^{-1}C)$ 是自由度为 k 的非中心卡方分布,

$$C=E[\mathrm{diag}(\omega_t)V_t^{-\frac{1}{2}}s_t]-A^{\mathrm{T}}J^{-1}E\left(\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}s_t\right)$$

以及 $\Sigma=\Sigma_0+A^{\mathrm{T}}J^{-1}IJ^{-1}A-2A^{\mathrm{T}}J^{-1}A$.

该结果表明得分型检验 T_n^2 对整体备择 $(\gamma=0)$ 和以速度 $n^{-\gamma}\,(0<\gamma<\frac{1}{2})$ 收敛到零假设的局部备择具有近似 1 的检验功效. 同时也表明该检验能侦察到以 $n^{-\frac{1}{2}}$ 的速度收敛到零假设模型的备择模型, 这个速度在拟合失度检验中是最快的. 在计算其功效时, 可以通过卡方分布决定其 p 值. T_n^2 的渐近功效是 $2-\Phi_k(\lambda_{\frac{\alpha}{2}}-\Sigma^{-\frac{1}{2}}C)-\Phi_k(\lambda_{\frac{\alpha}{2}}+\Sigma^{-\frac{1}{2}}C)$, $\Phi_k(\cdot)$ 是 k 维标准正态分布分布函数, $\lambda_{\frac{\alpha}{2}}$ 表示该分布的 $\frac{\alpha}{2}$ 分位数. 根据文献 [13, 引理 6.2.1],

$2-\Phi_k(\lambda_{\frac{\alpha}{2}}-\Sigma^{-\frac{1}{2}}C)-\Phi_k(\lambda_{\frac{\alpha}{2}}+\Sigma^{-\frac{1}{2}}C)$ 是 $C^{\mathrm{T}}\Sigma^{-1}C$ 的单调函数. 因此, 可选择使 $C^{\mathrm{T}}\Sigma^{-1}C$ 尽可能大的权函数 ω_t 以得到最优的检验. 注意到权函数与偏离函数 s_t 的内在关系. 为简单起见, 偏离函数 s_t 本身经常可作为权函数. 相关讨论可参考文献 [2,16].

众所周知, 备择并不总是方向型的. 更常见的, 我们只知道备择在某一个模型类中. 根据逼近理论, 对一备择, 或许可以考虑用样条来逼近, 从而可以假定备择属于样条基的一个凸组合类中. 因此, 如何构造检验统计量处理这种情形也许更具重要性和现实意义.

考虑模型序列:

$$H_{1n}: X_{nt}=m_t+C_n\sum_{j=1}^{d}\lambda_j s_{(j)t}+\varepsilon_t,\quad 1\leqslant t\leqslant n, \tag{8}$$

其中 $\lambda_1,\ldots,\lambda_d\in\mathbb{R}$ 是未知参数, $s_{(j)t}$ 是已知的可能偏离函数. 零假设模型对应

$$H_0:\lambda_1=\cdots=\lambda_d=0.$$

如上所述, 我们选择与可能偏离相匹配的权函数, 比如 $\omega_{(j)t}=s_{(j)t}$. 下面针对这一假设检验问题: H_0:$\|\lambda\|=0$ vs. H_1: $\|\lambda\|\geqslant\lambda_0$, 考虑如何构造构造极大极小 (maximin) 检验, 其中 $\|\cdot\|$ 是合适的范数, $\lambda^{\mathrm{T}}=(\lambda_1,\ldots,\lambda_d)$. 类似方向备择情形, 根据可能的偏离逐个考虑相应得分型检验 $\hat{T}_{n0}^{(j)}$, Σ_j 是对应 $\omega_{(j)t}$, $j=1,\ldots,d$ 得分检验的渐近方差, Σ_{nj} 是其经验形式. 令

$$\hat{T}_n=((\Sigma_{n1}^{-\frac{1}{2}}\hat{T}_{n0}^{(1)})^{\mathrm{T}},\ldots,(\Sigma_{nd}^{-\frac{1}{2}}\hat{T}_{n0}^{(d)})^{\mathrm{T}})^{\mathrm{T}}.$$

根据定理 2 的证明, 在备择模型 (8) 下, 当 $n\to\infty$, $n^{\gamma}C_n\to r$, $r\neq 0$, $0\leqslant\gamma<\frac{1}{2}$, 则依概率有 $\hat{T}_n\longrightarrow\infty$, 当 $n^{\frac{1}{2}}C_n\to 1$ 时, 有 $\hat{T}_n\xrightarrow{D}\Sigma^{(1)}\lambda+N_{kd}(0,\Sigma^{(2)})$, 其中 $\Sigma^{(1)}=(\Sigma_{(ij)}^{(1)})_{1\leqslant i,j\leqslant d}$, $\Sigma^{(2)}=(\Sigma_{(ij)}^{(2)})_{1\leqslant i,j\leqslant d}$,

$$\Sigma_{(ij)}^{(1)}=\Sigma_i^{-\frac{1}{2}}\left\{E[\mathrm{diag}(\omega_{(i)t})V_t^{-\frac{1}{2}}s_{(j)t}]-A_i^{\mathrm{T}}J^{-1}E\left(\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}s_{(j)t}\right)\right\},$$

$$\Sigma_{(ij)}^{(2)}=\Sigma_i^{-\frac{1}{2}}[E(\mathrm{diag}(\omega_{(i)t})\mathrm{diag}(\omega_{(j)t}))+A_i^{\mathrm{T}}J^{-1}IJ^{-1}A_j-A_i^{\mathrm{T}}J^{-1}A_j-A_j^{\mathrm{T}}J^{-1}A_i]\Sigma_j^{-\frac{1}{2}},$$

N_{kd} 表示 $\mathbb{R}^{kd}$ 上标准正态分布, $A_i=E[\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-\frac{1}{2}}\mathrm{diag}(\omega_{(i)t})], i=1,\ldots,d$.

根据 maximin 理论可构造得到最优的检验, 参考 [17, 定理 30.2]. 为此定义 $\Sigma_n^{(2)}=(\sigma_{(ij)n}^{(2)})_{1\leqslant i,j\leqslant d}$,

$$\Sigma_{(ij)n}^{(2)}=\frac{1}{n}\sum\Sigma_{ni}^{-\frac{1}{2}}[\mathrm{diag}(\hat{\omega}_{(i)t})\mathrm{diag}(\hat{\omega}_{(j)t})+\hat{A}_i^{\mathrm{T}}\hat{J}^{-1}\hat{I}\hat{J}^{-1}\hat{A}_j-\hat{A}_i^{\mathrm{T}}\hat{J}^{-1}\hat{A}_j-\hat{A}_j^{\mathrm{T}}\hat{J}^{-1}\hat{A}_i]\Sigma_{nj}^{-\frac{1}{2}}.$$

定理 3 对一给定显著性水平 α $(0<\alpha<1)$, $\gamma=\frac{1}{2}$, 检验

$$t=1_{\{\hat{T}_n^{\mathrm{T}}(\Sigma_n^{(2)})^{-1}\hat{T}_n\geqslant c_\alpha\}}$$

是关于 H_0 vs. H_1 假设检验问题的 maximin α 检验. H_1 是模型 (8), 并且 λ 满足 $\lambda^{\mathrm{T}}\Sigma^{(1)\mathrm{T}}\times(\Sigma^{(2)})^{-1}\Sigma^{(1)}\lambda\geqslant a$, 对任何用户给定的 $a>0$, c_α 是卡方分布 χ_{kd}^2 的 $1-\alpha$ 分位数. 渐近的 maximin 检验功效为 $P(\chi_{kd}^2(a)\geqslant c_\alpha)$, 其中 a 为非中心参数.

注记 1 根据定理 3, 对任意 d, 可得一个极大极小 (maximin) 检验. 注意到当 d 比较大时, χ_d^2 分布渐近等价于均值为 d 的正态分布. 用正态分布决定临界值, 需要正则化以得到标准的正态分布. 而这个正则化会降低检验的功效行为. 更多细节可参考文献 [18], 在该文中, 作者考虑了用零模型和备择模型拟合的平方距离的一个积分来构造检验统计量. 因此, 本文只需考虑一个适中的 d. 也可以借助非正式的残差图找到可能的偏离, 从而构造功效高的检验, 详细见第 4 节.

注记 2 假如备择是饱和备择情形, 有两个处理办法. 一个是利用逼近理论. 注意到任意 F_{t-1} 光滑函数可写成求和形式 $\sum_{j=1}^{\infty} c_j b_{jt}$, 其中 b_{jt} 是给定的基函数. 因此, 可以用 $\sum_{j=1}^{m} c_j b_{jt}$ 作为 s_t 的一个逼近, 其中 m 是用户选定的一个数. 并由此可构造一个 maximin 检验检查是否所有的 c_j 均为零. 这个逼近方法又反过来说明 maximin 检验的有用性. 另一个是构造万能检验. 如文献 [2] 所述, 可以在一个权函数类中考虑得分型检验的最大值作为检验统计量. 关于这一点, 第 5 节详细讨论.

3 非参数 Monte Carlo 检验步骤

注意到检验 T_n^2 涉及插入 (plug-in) 估计问题. 当样本量很小时, 极限分布不能起很好的作用. 在古典回归模型中, NMCT 方法常常被用来避免插入估计问题, 从而提高检验在小样本情形的功效, 相关的工作可参考文献 [6, 11–13]. 在相依数据情形, 如何推广 NMCT 方法还没有相关研究. 为此, 本节考虑把 NMCT 方法推广到时间序列模型.

在零假设 (1) 下, (5) 式成立, 也就是说, $T_{n0} = \frac{1}{\sqrt{n}} \sum (\eta_t \bullet \omega_t - A^{\mathrm{T}} J^{-1} \frac{\partial l_t}{\partial \theta}) + o_p(1)$. 然而在局部备择下, 比如模型 (7) 中 $C_n = \frac{1}{\sqrt{n}}$ 时, 有 $T_{n0} = C + \frac{1}{\sqrt{n}} \sum (\eta_t \bullet \omega_t - A^{\mathrm{T}} J^{-1} \frac{\partial l_t}{\partial \theta}) + o_p(1)$ (见 (24) 式). 为简单起见, 记 $g_t = \eta_t \bullet \omega_t - A^{\mathrm{T}} J^{-1} \frac{\partial l_t}{\partial \theta}$, $\hat{g}_t$ 表示 g_t 的估计. 下面给出相应的 NMCT 步骤.

步骤 1 产生独立同分布的随机变量 $e_t, t = 1, \ldots, n$, 均值为零、方差为 1, 且具有有界的支撑. 记 $E_n := \{e_t, t = 1, \ldots, n\}$, 定义 T_{n0} 条件副本 (conditional counterpart)

$$\tilde{T}_{n0}(E_n) = \frac{1}{\sqrt{n}} \sum \hat{g}_t e_t, \tag{9}$$

和 $T_n^{2*} = T_{n0}^{\mathrm{T}} T_{n0}$ 的条件副本 (conditional counterpart)

$$T_n^{2*}(E_n) = \left[\tilde{T}_{n0}(E_n)\right]^{\mathrm{T}} \tilde{T}_{n0}(E_n). \tag{10}$$

步骤 2 重复实验 u 次产生 u 个 E_n 的集合, 比如 $E_n^{(i)}, i = 1, \ldots, u$, 得到 u 个 $T_n^{2*}(E_n)$ 的值, 分别记为 $(T_n^{2*}(E_n))^{(i)}, i = 1, \ldots, u$.

步骤 3 记大于等于 T_n^{2*} 的 $(T_n^{2*}(E_n))^{(i)}$ 的个数为 v, 检验的 p-值的估计为 $\hat{p} = v/(u+1)$, 若 $\hat{p} \leqslant \alpha$ (α 为事先给定的水平), 则拒绝零假设.

注记 3 根据上面 3 个步骤, 可知这个 Monte Carlo 检验统计量 T_n^{2*} 不同于本文第 2 节构造的检验统计量 T_n^2. T_n^2 是尺度不变的, 而 T_n^{2*} 不是. 正是为了要达到尺度不变性, T_n^2 涉及插入 (plug-in) 估计, 从而导致在样本量小时检验功效的降低. 而 Monte Carlo 检验方法不需要插入 (plug-in) 估计, 因为相应的常数项抵消了.

定理 4 在零假设模型 (1) 及定理 1 条件下, 或在备择假设 (7) (当 $n \to \infty$ 时, $C_n \to 0$) 以及定理 2 的条件下, $T_n^{2*}(E_n)$ 的条件分布收敛到 T_n^{2*} 的渐近极限分布.

4 模拟和实际应用

为了验证得分型检验的有用性和可操作性, 本节进行一些 Monte Carlo 模拟实验和实际数据分析. 考虑 4 个二元时间序列自回归模型

$$M_1: \ X_t = \varepsilon_t,$$

$$M_2: \ X_t = a_1 X_{t-1} + \varepsilon_t,$$

$$M_3:\ X_t=\sin(a_2X_{t-1})+\varepsilon_t,$$
$$M_4:\ X_t=a_3X_{t-1}+a_4X_{t-2}+\varepsilon_t,$$

其中 $X_t=(X_{1t},X_{2t})^{\mathrm{T}},\varepsilon_t=(\varepsilon_{1t},\varepsilon_{2t})^{\mathrm{T}}$, 条件方差协方差为

$$V_t=\begin{pmatrix}0.2+0.2\varepsilon_{1,t-1}^2 & 0.1[(1+\varepsilon_{1,t-1}^2)(1+2\varepsilon_{2,t-1}^2)]^{\frac{1}{2}}\\ 0.1[(1+\varepsilon_{1,t-1}^2)(1+2\varepsilon_{2,t-1}^2)]^{\frac{1}{2}} & 0.2+0.4\varepsilon_{2,t-1}^2\end{pmatrix}. \tag{11}$$

上述四个模型中的误差项的条件异方差模型都是常相关系数的多元 GARCH 模型 (CC-MARCH/CC-MGARCH)(见文献 [19]). 根据前面的讨论, 为简单起见选择偏离函数作为权函数 ω_t. 在第一个实验中, M_1 是零假设模型, M_2 是备择模型. 参数值为 $a_1=0.1,0.2,0.3$, 样本量分别为 100、200 和 300. 为研究检验统计量对误差分布的稳健性, 考虑两种误差分布: (i) 正态分布; (ii) 自由度为 8 的标准化的学生氏分布 t_8. 权函数取为 X_{t-1}. 第二个实验与第一个的区别在于备择模型换为 M_3, 此时权函数选为 $\sin(X_{t-1})$. 接下来, 考虑零模型为 M_2, 备择模型为 M_4 之情形. 令 $a_1=a_3=0.1,a_4=0.1,0.2,0.3$, 样本量分别为 100、200 和 300. 为简单起见, 本实验仅仅考虑误差分布为正态分布的情形. 权函数选为 X_{t-2}. 上述实验重复 1 000 次得到检验的经验水平和经验功效. 模拟结果见表 1.

表 1 对二元时间序列模型检验统计量 T_n^2 的经验水平和经验功效. 名义水平为 5%

误差分布 (η_t)	数据产生过程	零模型	n =100	n=200	n=300
正态分布	M_1	M_1	0.057	0.046	0.051
	$M_2(a_1=0.1)$	M_1	0.118	0.225	0.365
	$M_2(a_1=0.2)$	M_1	0.439	0.777	0.931
	$M_2(a_1=0.3)$	M_1	0.833	0.993	1.000
	$M_3(a_2=0.1)$	M_1	0.134	0.256	0.377
	$M_3(a_2=0.2)$	M_1	0.471	0.772	0.925
	$M_3(a_2=0.3)$	M_1	0.828	0.991	1.000
	$M_2(a_1=0.1)$	M_2	0.048	0.050	0.051
	$M_4(a_4=0.1)$	M_2	0.133	0.271	0.424
	$M_4(a_4=0.2)$	M_2	0.548	0.881	0.975
	$M_4(a_4=0.3)$	M_2	0.859	0.975	0.985
t 分布	M_1	M_1	0.031	0.035	0.040
	$M_2(a_1=0.1)$	M_1	0.091	0.214	0.322
	$M_2(a_1=0.2)$	M_1	0.398	0.768	0.901
	$M_2(a_1=0.3)$	M_1	0.785	0.984	0.998
	$M_3(a_2=0.1)$	M_1	0.119	0.212	0.362
	$M_3(a_2=0.2)$	M_1	0.435	0.756	0.908
	$M_3(a_2=0.3)$	M_1	0.771	0.980	1.000

我们对极大极小 (maximin) 检验也进行一个模拟实验. 考虑模型 M_4, 让 $(a_3,a_4)=0$ 对应零模型, $(a_3,a_4)\neq 0$ 对应备择模型. 样本量分别为 100、200 和 300. 分别选 X_{t-1} 和 X_{t-2} 作为权函数 ω_{1t} 和 ω_{2t}. 为了表明极大极小 (maximin) 检验对误差分布 (新禧) 的稳健性, 分

别考虑了误差分布为正态和学生氏 t (比如标准化的学生氏分布 t_8) 分布两种情形. 上述实验重复 1 000 次得到检验的经验水平和经验功效. 模拟结果见表 2.

表 2　二元时间序列 M_4 模型极大极小(maximin)检验统计量的经验水平和经验功效. 名义水平为 5%

误差分布 (η_t)	数据产生过程	零模型	n =100	n=200	n=300
正态分布	$M_4(a_3=0.0,a_4=0.0)$	M_1	0.048	0.041	0.043
	$M_4(a_3=0.1,a_4=0.0)$	M_1	0.102	0.169	0.270
	$M_4(a_3=0.0,a_4=0.1)$	M_1	0.112	0.245	0.365
	$M_4(a_3=0.1,a_4=0.1)$	M_1	0.189	0.478	0.624
	$M_4(a_3=0.1,a_4=0.2)$	M_1	0.605	0.923	0.990
	$M_4(a_3=0.2,a_4=0.1)$	M_1	0.504	0.868	0.969
	$M_4(a_3=0.1,a_4=0.3)$	M_1	0.918	1.000	1.000
	$M_4(a_3=0.2,a_4=0.2)$	M_1	0.822	0.989	1.000
	$M_4(a_3=0.3,a_4=0.1)$	M_1	0.843	0.996	1.000
t 分布	$M_4(a_3=0.0,a_4=0.0)$	M_1	0.041	0.049	0.045
	$M_4(a_3=0.1,a_4=0.0)$	M_1	0.083	0.172	0.194
	$M_4(a_3=0.0,a_4=0.1)$	M_1	0.117	0.229	0.365
	$M_4(a_3=0.1,a_4=0.1)$	M_1	0.210	0.436	0.643
	$M_4(a_3=0.1,a_4=0.2)$	M_1	0.564	0.903	0.990
	$M_4(a_3=0.2,a_4=0.1)$	M_1	0.485	0.862	0.966
	$M_4(a_3=0.1,a_4=0.3)$	M_1	0.916	1.000	1.000
	$M_4(a_3=0.2,a_4=0.2)$	M_1	0.831	0.991	1.000
	$M_4(a_3=0.3,a_4=0.1)$	M_1	0.815	0.984	1.000

为对本文得分型检验和文献中存在的一些拟合优度检验比如混成 (portmanteau) 检验作一比较, 我们进行了一些模拟实验. 为简单起见, 考虑如下两个简单的向量自回归模型:

$$X_t = cX_{t-1} + \varepsilon_t, \tag{12}$$

$$X_t = cX_{t-1}^2 + \varepsilon_t, \tag{13}$$

其中 X_t 是二元时间序列, ε_t 服从二元标准正态分布. 对模型 (12) 和 (13), 我们分别研究了混成 (portmanteau) 检验 $Q(m), \tilde{Q}(m), Q^*(m)$ (见专著 [5, pp. 24–25]) 和得分型检验 T_n^2 的功效行为. $c=0$ 对应零假设, $c\neq 0$ 对应备择假设. 在得分型检验的构造中, 权函数分别选取为 X_{t-1} 和 X_{t-1}^2. 结果表明对如模型 (12) 的线性情形, 混成 (portmanteau) 检验表现比较理想, 但是我们的检验仍然要稍微好于它们. 而对于如 (13) 的非线性情形, 这些混成 (portmanteau) 检验表现很差, 我们的得分检验表现仍然很理想. 更多细节可见表 3 和 4.

当样本量不是太小, 比如 100、200 或 300 时, 用极限分布定临界值已经比较理想了. 当样本量比较小时, 可用非参数 Monte Carlo 检验 (NMCT) 方法. 本文一方面通过极限分布算得临界值, 另一方面通过 NMCT 方法算得临界值, 并且对两者得到的功效行为进行比较.

下面考察样本量比较小的两个例子.

例 1　考虑下面模型

$$X_t = cX_{t-1} + \varepsilon_t, \tag{14}$$

表 3 对模型 (12) 检验 $Q(m), \tilde{Q}(m)$, $Q^*(m)$, $m=1,2$ 和得分型检验 T_n^2 的经验水平和功效. 名义水平为 5%

样本量	检验	$c=0.0$ (size)	$c=0.1$ (power)	$c=0.2$ (power)	$c=0.3$ (power)	$c=0.4$ (power)	$c=0.5$ (power)	$c=0.6$ (power)
$n=50$	T_n^2	0.042	0.125	0.403	0.753	0.957	0.993	1.000
	$Q(1)$	0.076	0.071	0.201	0.483	0.802	0.950	0.998
	$\tilde{Q}(1)$	0.079	0.082	0.209	0.498	0.810	0.953	0.998
	$Q^*(1)$	0.077	0.076	0.204	0.490	0.806	0.952	0.998
	$Q(2)$	0.070	0.077	0.169	0.364	0.677	0.897	0.987
	$\tilde{Q}(2)$	0.077	0.085	0.180	0.386	0.692	0.907	0.989
	$Q^*(2)$	0.073	0.083	0.173	0.376	0.684	0.906	0.989
$n=100$	T_n^2	0.045	0.248	0.718	0.974	0.977	1.000	1.000
	$Q(1)$	0.063	0.154	0.494	0.896	0.990	1.000	1.000
	$\tilde{Q}(1)$	0.064	0.159	0.496	0.897	0.991	1.000	1.000
	$Q^*(1)$	0.063	0.155	0.495	0.897	0.990	1.000	1.000
	$Q(2)$	0.075	0.129	0.380	0.809	0.977	1.000	1.000
	$\tilde{Q}(2)$	0.076	0.133	0.387	0.817	0.980	1.000	1.000
	$Q^*(2)$	0.076	0.131	0.384	0.813	0.978	1.000	1.000

表 4 对模型 (13) 检验 $Q(m), \tilde{Q}(m)$, $Q^*(m)$, $m=1,2$ 和得分型检验 T_n^2 的经验水平和功效. 名义水平为 5%

样本量	检验	$c=0.0$ (size)	$c=0.04$ (power)	$c=0.08$ (power)	$c=0.12$ (power)	$c=0.16$ (power)
$n=50$	T_n^2	0.044	0.071	0.217	0.423	0.669
	$Q(1)$	0.077	0.074	0.077	0.073	0.084
	$\tilde{Q}(1)$	0.081	0.081	0.082	0.080	0.090
	$Q^*(1)$	0.079	0.074	0.079	0.075	0.085
	$Q(2)$	0.068	0.090	0.066	0.072	0.079
	$\tilde{Q}(2)$	0.072	0.099	0.073	0.082	0.085
	$Q^*(2)$	0.070	0.093	0.070	0.077	0.084
$n=100$	T_n^2	0.041	0.124	0.404	0.759	0.927
	$Q(1)$	0.077	0.073	0.076	0.082	0.114
	$\tilde{Q}(1)$	0.079	0.076	0.078	0.083	0.117
	$Q^*(1)$	0.078	0.074	0.077	0.082	0.114
	$Q(2)$	0.086	0.073	0.088	0.081	0.113
	$\tilde{Q}(2)$	0.089	0.078	0.096	0.085	0.117
	$Q^*(2)$	0.088	0.077	0.093	0.083	0.114

其中 X_t 是二变量时间序列, 误差项 ε_t 满足 $E(\varepsilon_t|X_{t-1})=0$, 条件方差协方差 V_t 为 (11). $c=0$ 对应零假设模型, $c\neq 0$ 对应备择模型. 样本量分别取为 20、40 和 60. 我们用 $c=0.1,0.2,\ldots,1.0$ 研究检验的功效. 权函数选为偏离函数 X_{t-1}. 重复上述实验 1 000 次得到检验的经验水平和经验功效.

例 2 考虑下面模型

$$X_{1,t} = \sin(cX_{1,t-1}) + \varepsilon_{1,t}, \tag{15}$$

$$X_{2,t} = \sin(cX_{2,t-1}) + \varepsilon_{2,t}, \tag{16}$$

其中 $X_t = (X_{1,t}, X_{2,t})^{\mathrm{T}}$ 是二变量时间序列, 误差项 $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t})^{\mathrm{T}}$ 满足 $E(\varepsilon_t|X_{t-1}) = 0$, 条件方差协方差 V_t 为 (11). $c = 0$ 对应零模型, $c \neq 0$ 对应备择模型. 样本量分别取为 20、40 和 60. 用 $c = 0.1, 0.2, \ldots, 1.0$ 研究检验的功效. 用偏离函数的一个近似作为权函数 $\omega_t^{\mathrm{T}} = (\sin(X_{1,t-1}), \sin(X_{2,t-1}))$. 重复上述实验 1000 次得到检验的经验水平和经验功效.

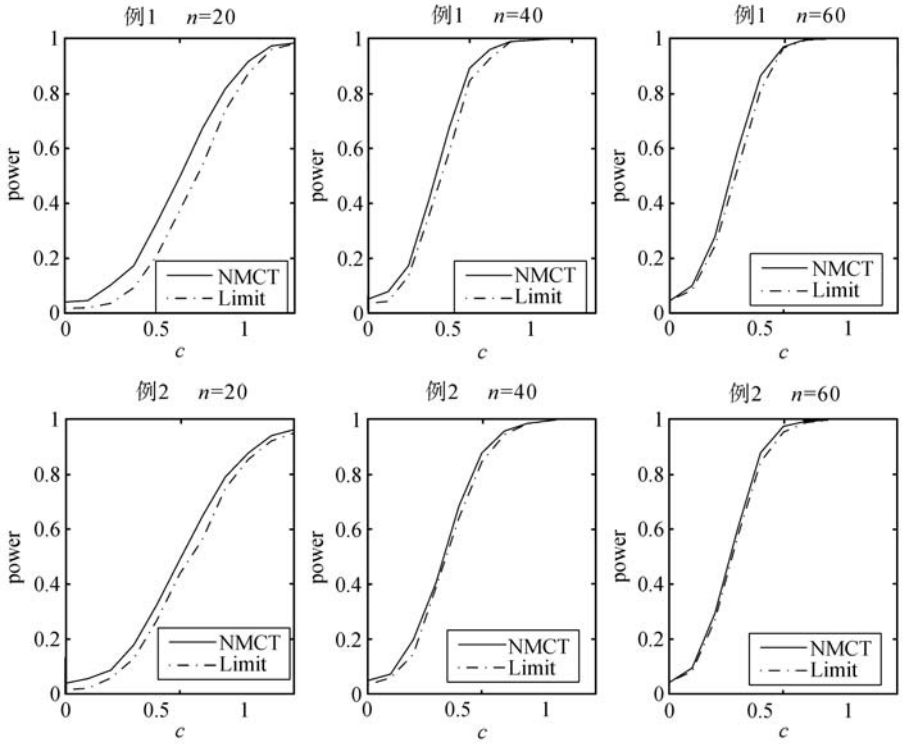

图 1 本图是例 1 和例 2 中检验统计量关于偏离系数 c 的功效图

上图结果表明当样本量小的时候 NMCT 方法更具有优势.

下面进行实际数据分析. 数据是香港和日本股票市场 1996 年 1 月 1 日到 1997 年 10 月 16 日的 469 个日 (对数) 收益率 (见文献 [1, 例 9.1, p. 365]). 样本选取时考虑到避免亚洲金融危机的影响. 记 r_{1t} 是香港指数收益率, r_{2t} 是日本指数收益率. Tsay[1] 指出这两个市场的波动率不是动态相关的, 但是两者是同时发生的 (contemporaneously) 或者说是同步相关的. 为了表明权函数的选择问题, 我们检验均值模型回归函数 $m_t = 0$ 是否成立? 若成立, 则这些实际数据来自一个简单 (均值模型回归函数为 0) 的条件异方差模型. 根据 Tsay[1] 分析结果, 可认为时间序列的方差模型为常相关系数多元 GARCH 型模型. 因此假设条件方差模型为如下常相关系数二元 GARCH 型模型

$$V_t = \begin{pmatrix} V_{t11} & V_{t12} \\ V_{t21} & V_{t22} \end{pmatrix},$$

$$V_{t11} = a_1 + a_2\varepsilon_{1,t-1}^2 + a_3 V_{t-1,11}, V_{t22} = a_4 + a_5\varepsilon_{2,t-1}^2 + a_6 V_{t-1,22},$$
$$V_{t12} = V_{t21} = a_7(V_{t11}V_{t22})^{\frac{1}{2}}.$$

根据未报道的非正式的残差图, 我们发现 r_{1t} 和 $r_{1,t-6}$ 的线性关系, 这与 Tsay[1] 的分析是一致的. 用 $(r_{1,t-6}, r_{2,t-6})$, $(1,1)$ (没有特别的挑选权函数) 作为两个权函数, 构造了极大极小 (maximin) 检验. 计算得检验的 p 值为 6.8044×10^{-11} $(P(\chi_4^2 > 53.467))$, 拒绝原假设. 这跟 Tsay[1] 的分析结果吻合.

5 总结和讨论

本文对时间序列形式一般的向量自回归模型的拟合优度检验问题提出一个新的方法. 构造的检验是得分型, 涉及权函数的选择. 当备择是方向型时, 可构造最优的得分检验. 然而当备择不是方向型时, 但在某个可能的备择类中时, 可基于得分型检验思想构造渐近分布自由的极大极小 (maximin) 检验. 由于这个备择类很大, 包含了大部分人们感兴趣的模型. 然而当我们对备择模型没有任何信息时, 也即所谓的饱和备择情形, 检验问题是完全非参数的, 需要构造一个万能 (omnibus) 检验来处理这个问题. 据我们所知, 相关领域还未见万能检验. 基于得分的极大极小检验结合非正式的残差图可部分地解决这个问题. 注意到本文的检验由于引入得分构造检验, 通过非正式的残差图可以帮助我们选择权函数构造高功效的检验. 这是该检验的一大优点, 尽管某种意义上增加了权函数的选择问题. 假如确实对备择没有任何信息, 即必须考虑所谓的饱和备择, 则可以考虑一个权函数类, 并相应构造一个基于过程的得分型检验. 具体如下. 设 $w(\cdot, s)$ 是一个权函数类, 指标 $s \in S$, 其中 S 是 $\mathbb{R}^k$ 的一个子集. 对每一权函数 $w(\cdot, s)$, 构造一个得分检验 $T_n(s)$. 最后把这些得分检验在指标集上的最大值作为检验 $\max_s |T_n(s)|$. 相关的讨论可参考文献 [2,9]. 然而, 付出的代价是这个问题因为涉及协变量的相依性很难研究, 极限分布也不再可用. 这个问题值得进一步研究.

当样本量很小时, 由于检验涉及插入 (plug-in) 估计从而导致检验功效低下, 我们发展非参数 Monte Carlo 方法决定临界值以避免使用插入 (plug-in) 估计导致的问题.

另外, Bauwens 等 [20] 指出, 多元时间序列 GARCH-型模型的诊断检验的发展是当前十个公开研究的问题之一, 并强调说这些问题任何一个问题的推进都将对多元 GARCH-型模型的理论研究和实际应用产生重要的的推动作用。幸运地是, 本文的方法可完全用于多元 GARCH-型模型的模型检验. 具体而言, 假若检验如下多元 GARCH-型模型

$$\varepsilon_t\varepsilon_t^{\mathrm{T}} = V_t + \epsilon_t,$$

其中 $E(\epsilon_t|F_{t-1}) = 0$ (假定均值模型 $X_t = m_t + \varepsilon_t$ 是充分的, 而只考虑条件方差模型是否是充分的). 实际上, 通过拉直运算, 检验下面模型的充分性即可:

$$\mathrm{vech}(\varepsilon_t\varepsilon_t^{\mathrm{T}}) = \mathrm{vech}(V_t) + \mathrm{vech}(\epsilon_t),$$

其中 "vech(A)" 表示对对称 $m \times m$ 方阵 A 去掉上三角元素然后拉直得到的 $\frac{m(m+1)}{2}$ 维向量.

6 附录

首先给出一些假设条件

A.1 假设文献 [14, p. 167] 中的条件 A.1 满足.

A.2 权函数 ω_t 关于参数 θ 是连续可微的, $E(\eta_{ht}\eta_{it}\eta_{jt})=0,\ h,i,j=1,\ldots,k$, 存在 $k\times k$ 对角阵 Σ_0 和 $p\times k$ 矩阵 A, 使得下式成立,

$$\Sigma_0=E[\mathrm{diag}(\omega_t)\mathrm{diag}(\omega_t)] \text{ 和 } A=E\left[\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-\frac{1}{2}}\mathrm{diag}(\omega_t)\right]. \tag{17}$$

A.3 存在真值 θ 的一个开邻域 $\mathcal{B}$, 使得

$$E\left[\sup_{\theta\in\mathcal{B}}\sum_{i=1}^{k}\left\|\frac{\partial^2(\eta_{it}\omega_{it})}{\partial\theta\partial\theta^{\mathrm{T}}}\right\|\right]<+\infty.$$

A.4 对所有的参数 $\dot\theta\in\Theta$ (Θ 是参数集), $J(\dot\theta):=-E(\frac{1}{n}\sum\frac{\partial^2 l_t}{\partial\theta\partial\theta^{\mathrm{T}}}|_{\theta=\dot\theta})$ 是一个正定矩阵.

A.5 存在真值 θ 一个开邻域 $\mathcal{B}$, 使得

$$E\left(\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}s_t\right)<\infty,\quad E\left(\frac{\partial[\mathrm{vec}(V_t)]^{\mathrm{T}}}{\partial\theta}[(V_t^{-1}s_t)\otimes V_t^{-1}]\right)<\infty,$$

$$E\left(\frac{\partial[\mathrm{vec}(V_t)]^{\mathrm{T}}}{\partial\theta}[V_t^{-\frac{1}{2}}\otimes(V_t^{-1}s_t)]\right)<\infty,\quad E(\mathrm{diag}(s_t)V_t^{-\frac{1}{2}}s_t)<\infty,$$

$$E(M_{1t})<\infty,\quad E(M_{2t})<\infty,\quad E\left[\frac{\partial[\mathrm{vec}(V_t)]^{\mathrm{T}}}{\partial\theta}((V_t^{-1}s_t)^{\mathrm{T}}\otimes(V_t^{-1}s_t)^{\mathrm{T}})\right]<\infty,$$

其中 M_{1t} 和 M_{2t} 分别来自 (19)–(20) 式.

定理 1 的证明 在零假设 (1) 下, 根据条件 A.1–A.3 以及 Taylor 展开可得

$$\begin{aligned}T_{n0}&=\frac{1}{\sqrt{n}}\sum\hat\eta_t\bullet\hat\omega_t=\frac{1}{\sqrt{n}}\sum\eta_t\bullet\omega_t+\left(\frac{1}{n}\sum\frac{\partial(\eta_t\bullet\omega_t)}{\partial\theta}\right)^{\mathrm{T}}\sqrt{n}(\hat\theta-\theta)\\&\quad+\frac{1}{2}\left[\sqrt{n}(\hat\theta-\theta)^{\mathrm{T}}n^{-\frac{3}{2}}\sum\frac{\partial^2(\eta_{1t}\omega_{1t})}{\partial\theta\partial\theta^{\mathrm{T}}}\bigg|_{\theta=\tilde\theta}\sqrt{n}(\hat\theta-\theta),\ldots,\right.\\&\qquad\left.\sqrt{n}(\hat\theta-\theta)^{\mathrm{T}}n^{-\frac{3}{2}}\sum\frac{\partial^2(\eta_{kt}\omega_{kt})}{\partial\theta\partial\theta^{\mathrm{T}}}\bigg|_{\theta=\tilde\theta}\sqrt{n}(\hat\theta-\theta)\right]^{\mathrm{T}}\\&=\frac{1}{\sqrt{n}}\sum\eta_t\bullet\omega_t+\left(\frac{1}{n}\sum\frac{\partial(\eta_t\bullet\omega_t)}{\partial\theta}\right)^{\mathrm{T}}\sqrt{n}(\hat\theta-\theta)+o_p(1)\\&=\frac{1}{\sqrt{n}}\sum(\eta_t\bullet\omega_t-A^{\mathrm{T}}J^{-1}\frac{\partial l_t}{\partial\theta})+o_p(1),\end{aligned}\tag{18}$$

其中 $\tilde\theta$ 是界于 $\hat\theta$ 和真值 θ 之间的一个向量.

根据鞅中心极限定理 (见文献 [15, p. 788]) 和 Cramér-Wold 技术, 可得 $T_{n0}\xrightarrow{D}N(0,\Sigma)$, 其中

$$\Sigma=\mathrm{var}\left(\eta_t\bullet\omega_t-A^{\mathrm{T}}J^{-1}\frac{\partial l_t}{\partial\theta}\right)=E[\mathrm{diag}(\omega_t)\mathrm{diag}(\omega_t)]+A^{\mathrm{T}}J^{-1}IJ^{-1}A-2A^{\mathrm{T}}J^{-1}A.$$

于是有

$$T_{n0}^{\mathrm{T}}\Sigma^{-1}T_{n0}\xrightarrow{D}\chi_k^2.$$

注意到 (6) 式以及 Σ_n 在零假设下是 Σ 的相合估计这一事实, 并根据 Slutsky 定理可得

$$T_n^2=T_{n0}^{\mathrm{T}}\Sigma_n^{-1}T_{n0}\xrightarrow{D}\chi_k^2.$$

为研究检验的功效, 下面给出一个引理以简化主要定理的证明.

引理 1 在备择假设 (7) 和条件 A.1–A.5 下, $C_n=n^{-\gamma}$, $0\leqslant\gamma\leqslant\frac{1}{2}$, 拟极大似然估计 (QMLE) $\hat\theta$ 不是根号 n 相合的.

证明 不失一般性, 我们考虑情形 $0<\gamma\leqslant\frac{1}{2}$.

在备择假设 H_{1n} (7) 下, 可得

$$\begin{aligned}\frac{\partial l_t}{\partial\theta}=&\frac{1}{2}\frac{\partial[\mathrm{vec}(V_t)]^{\mathrm{T}}}{\partial\theta}\mathrm{vec}[V_t^{-1}(X_{nt}-m_t)(X_{nt}-m_t)^{\mathrm{T}}V_t^{-1}-V_t^{-1}]+\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}(X_{nt}-m_t)\\
=&\frac{1}{2}\frac{\partial[\mathrm{vec}(V_t)]^{\mathrm{T}}}{\partial\theta}\mathrm{vec}[V_t^{-1}(X_{nt}-m_t-C_ns_t)(X_{nt}-m_t-C_ns_t)^{\mathrm{T}}V_t^{-1}-V_t^{-1}]+\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}\\
&\times(X_{nt}-m_t-C_ns_t)+C_n\bigg\{\frac{1}{2}\frac{\partial[\mathrm{vec}(V_t)]^{\mathrm{T}}}{\partial\theta}\mathrm{vec}[V_t^{-1}s_t(X_{nt}-m_t-C_ns_t)^{\mathrm{T}}V_t^{-1}\\
&+V_t^{-1}(X_{nt}-m_t-C_ns_t)s_t^{\mathrm{T}}V_t^{-1}+C_nV_t^{-1}s_ts_t^{\mathrm{T}}V_t^{-1}]+\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}s_t\bigg\}\\
=&\frac{1}{2}\frac{\partial[\mathrm{vec}(V_t)]^{\mathrm{T}}}{\partial\theta}\mathrm{vec}[V_t^{-\frac{1}{2}}\eta_t\eta_t^{\mathrm{T}}V_t^{-\frac{1}{2}}-V_t^{-1}]+\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-\frac{1}{2}}\eta_t\\
&+C_n\bigg\{\frac{1}{2}\frac{\partial[\mathrm{vec}(V_t)]^{\mathrm{T}}}{\partial\theta}\mathrm{vec}[V_t^{-1}s_t\eta_t^{\mathrm{T}}V_t^{-\frac{1}{2}}+V_t^{-1}\eta_ts_t^{\mathrm{T}}V_t^{-1}+C_nV_t^{-1}s_ts_t^{\mathrm{T}}V_t^{-1}]+\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}s_t\bigg\}.\end{aligned}$$

记 $\frac{\partial l_t^{(1)}}{\partial\theta}=\frac{1}{2}\frac{\partial[\mathrm{vec}(V_t)]^{\mathrm{T}}}{\partial\theta}\mathrm{vec}[V_t^{-\frac{1}{2}}\eta_t\eta_t^{\mathrm{T}}V_t^{-\frac{1}{2}}-V_t^{-1}]+\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-\frac{1}{2}}\eta_t$, $\frac{\partial l_t^{(2)}}{\partial\theta}=\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}s_t$.

根据条件 $A.5$, 可得

$$\begin{aligned}\frac{1}{n}\sum\frac{\partial l_t^{(1)}}{\partial\theta}&=\frac{1}{n}\sum\frac{\partial l_t^{(1)}}{\partial\theta}+C_n\frac{1}{n}\sum\frac{\partial l_t^{(2)}}{\partial\theta}+o_p(C_n),\\
\frac{1}{n}\sum\frac{\partial l_t^{(2)}}{\partial\theta}&=\frac{1}{n}\sum\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}s_t+O_p(C_n).\end{aligned}$$

对上面的函数 $\frac{\partial l_t}{\partial\theta}$ 微分并对 $F_{n,t-1}$ 取迭代的期望, 可得下面的 $p\times p$ 对称, 正定半正定的 Hesse 矩阵:

$$\begin{aligned}&\sum E\bigg(-\frac{\partial^2 l_t}{\partial\theta\partial\theta^{\mathrm{T}}}|F_{n,t-1}\bigg)\\
&\quad=\frac{1}{2}\sum\frac{\partial[\mathrm{vec}(V_t)]^{\mathrm{T}}}{\partial\theta}(V_t^{-1}\otimes V_t^{-1})\frac{\partial[\mathrm{vec}(V_t)]}{\partial\theta^{\mathrm{T}}}+\sum\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}\frac{\partial m_t}{\partial\theta^{\mathrm{T}}}-C_n\sum M_t\\
&\quad=:\sum J_t-C_n\sum M_{1t}-C_n^2\sum M_{2t},\end{aligned}$$

其中

$$\begin{aligned}J_t&=\frac{1}{2}\frac{\partial[\mathrm{vec}(V_t)]^{\mathrm{T}}}{\partial\theta}(V_t^{-1}\otimes V_t^{-1})\frac{\partial[\mathrm{vec}(V_t)]}{\partial\theta^{\mathrm{T}}}+\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}\frac{\partial m_t}{\partial\theta^{\mathrm{T}}},\\
M_{1t}&=(I_p\otimes s_t^{\mathrm{T}})\frac{\partial[\mathrm{vec}(V_t^{-1})]^{\mathrm{T}}}{\partial\theta}\frac{\partial m_t}{\partial\theta^{\mathrm{T}}}+(I_p\otimes(s_t^{\mathrm{T}}V_t^{-1}))\frac{\partial^2 m_t}{\partial\theta\partial\theta^{\mathrm{T}}}\\
&\quad-\frac{1}{2}\bigg(s_t^{\mathrm{T}}\otimes\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}+\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}\otimes s_t^{\mathrm{T}}\bigg)(V_t^{-1}\otimes V_t^{-1})\frac{\partial\mathrm{vec}(V_t)}{\partial\theta^{\mathrm{T}}},\qquad(19)\\
M_{2t}&=\frac{1}{2}[I_p\otimes((V_t^{-1}s_t)^{\mathrm{T}}\otimes(V_t^{-1}s_t)^{\mathrm{T}})]\frac{\partial^2\mathrm{vec}(V_t)}{\partial\theta\partial\theta^{\mathrm{T}}}\\
&\quad+\frac{1}{2}(I_p\otimes(s_t^{\mathrm{T}}\otimes s_t^{\mathrm{T}}))\frac{\partial(V_t^{-1}\otimes V_t^{-1})}{\partial\theta}\frac{\partial\mathrm{vec}(V_t)}{\partial\theta^{\mathrm{T}}}.\qquad(20)\end{aligned}$$

定义 $J=\frac{1}{n}\sum J_t$, $M_1=\frac{1}{n}\sum M_{1t}$ 和 $M_2=\frac{1}{n}\sum M_{2t}$. 因为 $\hat\theta$ 是使似然函数 l 极大的拟极大似然估计, 根据中值定理和条件 A.4, 可得

$$\hat\theta-\theta=\bigg[-\frac{1}{n}\sum\frac{\partial^2 l_t}{\partial\theta\partial\theta^{\mathrm{T}}}\bigg|_{\theta=\hat\theta}\bigg]^{-1}\frac{1}{n}\sum\frac{\partial l_t}{\partial\theta}$$

$$
\begin{aligned}
&= [J(\theta)+o_p(1)]^{-1}\frac{1}{n}\sum\frac{\partial l_t^{(1)}}{\partial\theta}+C_n[J(\theta)+o_p(1)]^{-1}\frac{1}{n}\sum\frac{\partial l_t^{(2)}}{\partial\theta}\\
&= J(\theta)^{-1}\frac{1}{n}\sum\frac{\partial l_t^{(1)}}{\partial\theta}+C_nJ(\theta)^{-1}\frac{1}{n}\sum\frac{\partial l_t^{(2)}}{\partial\theta},
\end{aligned}
$$

该参数估计是相合的. 注意到在零假设下, $\frac{1}{n}\sum\frac{\partial l_t}{\partial\theta}$ 依概率收敛到零, 这个收敛性导致参数估计 $\hat\theta$ 收敛到 θ. 更多细节可参考文献 [14]. 这一标准的做法同样适用我们的问题, 只要注意到在局部备择假设 (7), $\frac{1}{n}\sum\frac{\partial l_t}{\partial\theta}$ 也是依概率收敛到零, 并且 $\hat\theta-\theta=o_p(1)$. 进一步, 该参数估计在局部备择下不是根号 n 相合的.

$$
\begin{aligned}
n^{\frac{1}{2}}(\hat\theta-\theta) &= (J+o_p(1))^{-1}n^{-\frac{1}{2}}\sum\frac{\partial l_t^{(1)}}{\partial\theta}+C_n(J+o_p(1))^{-1}n^{-\frac{1}{2}}\sum\frac{\partial l_t^{(2)}}{\partial\theta}+o_p(C_nn^{\frac{1}{2}})\\
&= J^{-1}n^{-\frac{1}{2}}\sum\frac{\partial l_t^{(1)}}{\partial\theta}+C_nJ^{-1}n^{-\frac{1}{2}}\sum\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}s_t+o_p(C_nn^{\frac{1}{2}}),
\end{aligned}
$$

也即, 主要偏差项是 $C_nn^{\frac{1}{2}}J^{-1}\frac{1}{n}\sum\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}s_t$.

定理 2 的证明 首先考虑情形 $0<\gamma<\frac{1}{2}$. 加权的残差可分解如下:

$$
\begin{aligned}
T_{n0} &= \frac{1}{\sqrt{n}}\sum[\hat V_t^{-\frac{1}{2}}(X_{nt}-\hat m_t)]\bullet\hat\omega_t\\
&= \frac{1}{\sqrt{n}}\sum[V_t^{-\frac{1}{2}}(X_{nt}-m_t)]\bullet\omega_t+\frac{1}{n}\sum\frac{\partial([V_t^{-\frac{1}{2}}(X_{nt}-m_t)]\bullet\omega_t)}{\partial\theta^{\mathrm{T}}}\sqrt{n}(\hat\theta-\theta)\\
&\quad+\Bigg[\sqrt{n}(\hat\theta-\theta)^{\mathrm{T}}n^{-\frac{3}{2}}\sum\frac{\partial^2([I_{k(1)}^{\mathrm{T}}V_t^{-\frac{1}{2}}(X_{nt}-m_t)]\omega_{1t})}{\partial\theta\partial\theta^{\mathrm{T}}}\Bigg|_{\theta=\tilde{\tilde\theta}}\sqrt{n}(\hat\theta-\theta),\ldots,\\
&\quad\sqrt{n}(\hat\theta-\theta)^{\mathrm{T}}n^{-\frac{3}{2}}\sum\frac{\partial^2([I_{k(k)}^{\mathrm{T}}V_t^{-\frac{1}{2}}(X_{nt}-m_t)]\omega_{kt})}{\partial\theta\partial\theta^{\mathrm{T}}}\Bigg|_{\theta=\tilde{\tilde\theta}}\sqrt{n}(\hat\theta-\theta)\Bigg]^{\mathrm{T}}\\
&= \frac{1}{\sqrt{n}}\sum[V_t^{-\frac{1}{2}}(X_{nt}-m_t)]\bullet\omega_t+\frac{1}{n}\sum\frac{\partial[V_t^{-\frac{1}{2}}(X_{nt}-m_t)]\bullet\omega_t}{\partial\theta^{\mathrm{T}}}\sqrt{n}(\hat\theta-\theta)+o_p(1)\\
&= \frac{1}{\sqrt{n}}\sum[V_t^{-\frac{1}{2}}(X_{nt}-m_t-C_ns_t)]\bullet\omega_t+C_nn^{\frac{1}{2}}\frac{1}{n}\sum(V_t^{-\frac{1}{2}}s_t)\bullet\omega_t\\
&\quad+\Bigg(-A^{\mathrm{T}}+C_n\frac{1}{n}\sum(V_t^{-\frac{1}{2}}s_t)\bullet\frac{\partial\omega_t}{\partial\theta^{\mathrm{T}}}\Bigg)J^{-1}\frac{\partial l_t}{\partial\theta}+o_p(1)\\
&= \frac{1}{\sqrt{n}}\sum\Bigg\{[V_t^{-\frac{1}{2}}(X_{nt}-m_t-C_ns_t)]\bullet\omega_t-A^{\mathrm{T}}J^{-1}\frac{\partial l_t^{(1)}}{\partial\theta}\Bigg\}\\
&\quad+C_nn^{\frac{1}{2}}\frac{1}{n}\sum\Bigg[(V_t^{-\frac{1}{2}}s_t)\bullet\omega_t-A^{\mathrm{T}}J^{-1}\frac{\partial l_t^{(2)}}{\partial\theta}\Bigg]+o_p(C_nn^{\frac{1}{2}}),
\end{aligned}
\tag{21}
$$

其中 $\tilde{\tilde\theta}$ 是中值, $I_{k(1)},\ldots,I_{k(k)}$ 表示 I_k 的列向量. 容易看出 T_{n0} 的主项是 $C_nn^{\frac{1}{2}}\frac{1}{n}\sum[(V_t^{-\frac{1}{2}}s_t)\bullet\omega_t-A^{\mathrm{T}}J^{-1}\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}s_t]$.

当 $0<\gamma<\frac{1}{2}$, 通过选择满足如下不等式的权函数

$$
E\Bigg\{\Bigg[V_t^{-\frac{1}{2}}s_t-V_t^{-\frac{1}{2}}\frac{\partial m_t}{\partial\theta^{\mathrm{T}}}J^{-1}E\Bigg(\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}s_t\Bigg)\Bigg]\bullet\omega_t\Bigg\}\neq 0,
$$

检验统计量 T_{n0} 的主要偏差为

$$
C_nn^{\frac{1}{2}}\frac{1}{n}\sum\Bigg[(V_t^{-\frac{1}{2}}s_t)\bullet\omega_t-A^{\mathrm{T}}J^{-1}\frac{\partial m_t^{\mathrm{T}}}{\partial\theta}V_t^{-1}s_t\Bigg]
$$

$$
\begin{aligned}
&= C_n n^{\frac{1}{2}} E\left\{\left[V_t^{-\frac{1}{2}} s_t - V_t^{-\frac{1}{2}} \frac{\partial m_t}{\partial \theta^{\mathrm{T}}} J^{-1} E\left(\frac{\partial m_t^{\mathrm{T}}}{\partial \theta} V_t^{-1} s_t\right)\right] \bullet \omega_t\right\} + o_p(C_n n^{\frac{1}{2}}) \\
&\longrightarrow \infty.
\end{aligned}
$$

因此, 依概率有 $T_n \to \infty$. 当 $\gamma = \frac{1}{2}$, 则依概率有

$$
\begin{aligned}
&C_n n^{\frac{1}{2}} \frac{1}{n} \sum \left[(V_t^{-\frac{1}{2}} s_t) \bullet \omega_t - A^{\mathrm{T}} J^{-1} \frac{\partial m_t^{\mathrm{T}}}{\partial \theta} V_t^{-1} s_t\right] \\
&\qquad \longrightarrow E\left\{\left[V_t^{-\frac{1}{2}} s_t - V_t^{-\frac{1}{2}} \frac{\partial m_t}{\partial \theta^{\mathrm{T}}} J^{-1} E\left(\frac{\partial m_t^{\mathrm{T}}}{\partial \theta} V_t^{-1} s_t\right)\right] \bullet \omega_t\right\},
\end{aligned} \tag{22}
$$

和

$$
\begin{aligned}
&E\left[\left(V_t^{-\frac{1}{2}}\left(X_t - m_t - \frac{1}{\sqrt{n}} s_t\right)\right) \bullet \omega_t - A^{\mathrm{T}} J^{-1} \frac{\partial l_t^{(1)}}{\partial \theta}\right]\left[(V_t^{-\frac{1}{2}}(X_t - m_t - n^{-\frac{1}{2}} s_t)) \bullet \omega_t\right. \\
&\qquad \left. - A^{\mathrm{T}} J^{-1} \frac{\partial l_t^{(1)}}{\partial \theta}\right]^{\mathrm{T}} \\
&\qquad \longrightarrow E\left[(V_t^{-\frac{1}{2}}(X_t - m_t)) \bullet \omega_t - A^{\mathrm{T}} J^{-1} \frac{\partial l_t}{\partial \theta}\right]\left[(V_t^{-\frac{1}{2}}(X_t - m_t)) \bullet \omega_t - A^{\mathrm{T}} J^{-1} \frac{\partial l_t}{\partial \theta}\right]^{\mathrm{T}} \\
&\qquad = \Sigma,
\end{aligned}
$$

以及

$$
\frac{1}{\sqrt{n}} \sum \left\{\left[V_t^{-\frac{1}{2}}(X_{nt} - m_t - C_n s_t)\right] \bullet \omega_t - A^{\mathrm{T}} J^{-1} \frac{\partial l_t^{(1)}}{\partial \theta}\right\} \xrightarrow{D} N(0, \Sigma). \tag{23}
$$

根据 (21)–(23) 式, 可得

$$
T_{n0} = C + n^{-\frac{1}{2}} \sum \left(\eta_t \bullet \omega_t - A^{\mathrm{T}} J^{-1} \frac{\partial l_t}{\partial \theta}\right) + o_p(1), \tag{24}
$$

其中 $\eta_t = V_t^{-\frac{1}{2}}(X_t - m_t - n^{-\frac{1}{2}} s_t)$, $C = E\{[(V_t^{-\frac{1}{2}} s_t) - V_t^{-\frac{1}{2}} \frac{\partial m_t}{\partial \theta^{\mathrm{T}}} C_0] \bullet \omega_t\} = E[\mathrm{diag}(\omega_t) V_t^{-\frac{1}{2}} s_t] - A^{\mathrm{T}} J^{-1} E(\frac{\partial m_t^{\mathrm{T}}}{\partial \theta} V_t^{-1} s_t)$.

于是有

$$
T_{n0} \xrightarrow{D} N(C, \Sigma) \qquad \text{和} \qquad T_n^2 \xrightarrow{D} \chi_k^2(C^{\mathrm{T}} \Sigma^{-1} C),
$$

其中 $\Sigma = E(\mathrm{diag}(\omega_t)\mathrm{diag}(\omega_t)) + A^{\mathrm{T}} J^{-1} I J^{-1} A - 2A^{\mathrm{T}} J^{-1} A$, $C = E[(V_t^{-\frac{1}{2}} s_t) \bullet \omega_t - A^{\mathrm{T}} J^{-1} E(\frac{\partial m_t^{\mathrm{T}}}{\partial \theta} \times V_t^{-1} s_t)]$.

根据上面类似的讨论, 对整体备择, 也即 $\gamma = 0$ 情形, 也可得

$$
T_n^2 \longrightarrow \infty.
$$

定理 4 的证明 定理 4 的证明完全类似专著 [13] 中定理 5.3.1 或定理 6.3.2 的证明. 要证明这个定理, 只需证明 $\tilde{T}_{n0}(E_n)$ 的渐近正态性以及 $\tilde{T}_{n0}(E_n)$ 和 T_{n0} 的渐近协方差的一致性即可. 注意到 e_t 独立于 $g_t, t = 1, \ldots, n$. $g_t e_t$ 是鞅差序列, $\tilde{T}_{n0}(E_n)$ 协方差是 $\frac{1}{n} \sum_{t=1}^n \hat{g}_t \hat{g}_t^{\mathrm{T}}$. 根据参数估计的相合性, Taylor 展开以及弱大数定律, 易知 $\frac{1}{n} \sum_{t=1}^n \hat{g}_t \hat{g}_t^{\mathrm{T}}$ 收敛到 $E(g_t g_t^{\mathrm{T}})$, 而这正好是 T_{n0} 的极限协方差. 关于渐近正态性, 根据鞅中心极限定理可得其极限分布是 $N(0, E(g_t g_t^{\mathrm{T}}))$. 这完全等同于 T_{n0} 的极限分布.

致谢 感谢匿名审稿人具有建设性的意见和建议.

参考文献

1 Tsay R S. Analysis of Financial Time Series. New York: Wiley, 2002

2 Stute W, Zhu L X. Nonparametric checks for single-index models. *Ann Statist*, **33**: 1048–1083 (2005)

3 Wu J H, Zhu L X. A score type test for general autoregressive models in time series. *Acta Math Appl Sinica*, **23**(3): 439–450 (2007)

4 Wu J H, Zhu L X. Empirical likelihood ratio tests for multivariate regression models. *Front Math China*, **2**(1): 149–168 (2007)

5 Li W K. Diagnostic Checking in Time Series. Boca Raton: Chapman and Hall, 2004

6 He X, Zhu L X. A lack-of-fit test for quantile regression. *J Amer Statist Assoc*, **464**: 1013–1022 (2003)

7 Koul H L, Stute W. Nonparametric method checks for time series. *Ann Statist*, **27**: 204–236 (1999)

8 Stute W. Nonparametric model checks for regression. *Ann Statist,* **25**: 613–641 (1997)

9 Stute W, Thies S, Zhu L X. Model checks for regression: An innovation process approach. *Ann Statist*, **26**: 1916–1934 (1998)

10 Stute W, Zhu L X. Model checks for generalized linear models. *Scand J Statist*, **29**: 535–546 (2002)

11 Zhu L X. Model checking of dimension-reduction type for regression. *Statist Sinica*, **13**: 283–296 (2003)

12 Zhu L X, Ng K W. Checking the adequacy of a partial linear model. *Statist Sinica*, **13**: 763–781 (2003)

13 Zhu L X. Nonparametric Monte Carlo Tests and Their Applications. New York: Springer, 2005

14 Bollerslev T, Wooldridge J M. Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariance. *Econometric Rev*, **11**: 143–172 (1992)

15 Billingsley P. The Lindeberg-Levy theorem for martingales. *Proc Amer Math Soc*, **12**: 788–792 (1961)

16 Zhu L X, Cui H J. Testing the adequacy for a general linear error-in-variables model. *Statist Sinica*, **15**: 1049–1068 (2005)

17 Strasser H. Mathematical Theory of Statistics. Berlin: De Gruyter, 1985

18 Härdle W, Mammen E. Comparing non-parametric versus parametric regression fits. *Ann Statist*, **21**: 1926–1947 (1993)

19 Bollerslev T. Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Rev Econom Statist*, **72**: 498–505 (1990)

20 Bauwens L, Laurent S, Rombouts J. Multivariate GARCH models: A survey. *J Appl Econometrics*, **21**: 79–109 (2006)