

动态异方差随机前沿模型的Bayesian推断

程 迪¹, 张世斌²

(1. 内蒙古大学 数学科学学院, 内蒙古呼和浩特 010021;

2. 上海海事大学 数学系, 上海浦东 201306)

摘 要: 随机前沿模型中如果忽略单边干扰项的异质性(heterogeneity)往往导致错误的效率估计. 从个体特征的影响和方差的时变性两方面对单边干扰项进行考虑, 提出异方差动态随机前沿模型. 利用Gibbs抽样方法对动态异方差随机前沿模型进行Bayesian分析. 导出了模型参数的后验条件分布, 对中小样本的模拟实验显示在最小后验均方误差准则下得到的参数估计值非常接近真值. 对电力公司的实际数据进行分析显示对数无效率项的方差有一定的时变性.

关键词: 随机前沿模型; Bayesian分析; 异方差; Gibbs抽样; Metropolis-Hastings抽样

中图分类号: O212.8

文献标识码: A **文章编号:** 1000-4424(2016)02-0127-09

§1 引 言

随机前沿模型由Aigner等(1977)^[1]及Meeusen和Broeck(1977)^[2]提出, 被广泛用于衡量经济效率. 随机前沿模型将误差项分为对称误差项和单边干扰项两部分, 分别反映了测量误差和模型中的技术无效率对产出或成本造成的负面影响, 即单边干扰项又为无效率项. 通常假定无效率项在所有决策单元中是独立同分布的, 无效率项的分布一般假设为半正态分布^[1], 指数分布^[2], 截断正态分布^[3], Gamma分布^[4]或倒Gamma分布^[5]等. 然而, 影响一个公司效率的因素除了成本函数模型中的输入输出变量, 还包括一些公司特征, 如公司的所有权, 公司政策的改变, 行业的竞争等. 这些因素将对公司的效率产生一定的影响. 如果公司的这些特征产生异质性(heterogeneity)或者对单边干扰项产生影响, 这些特征就可以在前沿模型中表现出来. 文[6]将协变量直接以函数形式表现并研究了单边干扰项中未被观察到的异质性. 文[7]针对多输出随机前沿模型, 提出用copula函数来刻画无效率项分布. 文[8]提出了动态随机前沿模型, 无效率项随着时间的变化而变化, 并用Bayesian分析和MCMC方法对参数进行估计. 文[9]将随机前沿模型进行转换, 消除个体的固有影响, 并对其参数估计的一致性进行分析. 文[10]对无效率项服从广义Gamma分布和广义混合Gamma分布的随机前沿模型进行Bayesian分析, 考虑了无效率项中

收稿日期: 2015-11-13 修回日期: 2016-04-19

基金项目: 上海市自然科学基金(13ZR1419100); 上海市教委科研创新项目(14YZ115)

可观察到的异方差. 文[11]提出用参数来刻画未被观测到的异质性, 指出了异质性可能影响的变量并通过实例对变量进行分析论证.

现有随机前沿模型较少涉及无效率项中未观测到的异质性, 本文提出用GARCH(p, q)来刻画对数无效率项. 其中, 均值方程刻画可以观测到的异质性, 条件方差方程刻画未观测到的异质性. 当 $p = q = 0$ 时对数无效率项方差为常数, 无效率项不存在异方差. 否则, 对数无效率项方差是过去误差项和滞后条件方差的函数, 随时间变化方差的值不同. 本文利用Gibbs抽样对动态异方差随机前沿模型进行Bayesian分析. 通过对模型参数选取合理的先验分布, 导出模型参数的后验条件分布, 并给出了Gibbs抽样的具体策略. 由于参数 λ 和 α 的后验条件分布不是标准密度函数, 且极难找到标准分布的密度函数作为控制函数进行舍选, 常用的随机游走抽样和Griddy-Gibbs抽样方法在该分层随机前沿模型中的舍选效率非常低. 本文参照Nakatsuma(参见[12])提出的近似GARCH模型方法, 将条件方差方程构造成为新的ARMA方程, 得到参数的建议分布, 最后通过Metropolis-Hastings抽样实现对参数 λ 和 α 的抽样. 此方法大大提高了动态异方差随机前沿模型的计算效率.

§2 模型介绍

与截面数据相比, 面板数据更能提供对公司技术效率的可靠估计. 考虑随机前沿模型

$$y_{it} = \mathbf{x}_{it}'\boldsymbol{\beta} + v_{it} - u_{it}, \quad i = 1, 2, \dots, N, \quad t = 1, \dots, T.$$

其中 y_{it} 表示第 i 个个体 t 时间的产出, $\mathbf{x}_{it}$ 为 $k \times 1$ 维向量, 表示第 i 个个体 t 时间的投入, $\boldsymbol{\beta}$ 是 $k \times 1$ 的参数向量. v_{it} 是统计误差, 设 v_{it} 关于 i 和 t 相互独立并且有相同的分布 $v_{it} \sim N(0, \sigma^2)$. u_{it} 是无效率项, 为非负的随机变量. 假设 v_{it} 和 u_{it} 相互独立, 且无效率项 u_{it} 满足

$$\log u_{it} = \mathbf{z}_{it}'\boldsymbol{\gamma} + \varepsilon_{it}, \quad t = 1, 2, \dots, T, \quad (1)$$

$$\varepsilon_{it} = \sqrt{h_{it}}\xi_{it}, \quad h_{it} = \lambda_0 + \sum_{j=1}^p \lambda_j \varepsilon_{i,t-j}^2 + \sum_{j=1}^q \alpha_j h_{i,t-j}.$$

对所有的 i, t 都有 $\xi_{it} \sim N(0, 1)$ 且相互独立, $\mathbf{z}_{it}$ 为 $m \times 1$ 维的协变量, $\boldsymbol{\gamma}$ 为 $m \times 1$ 维的参数向量. 这里 $\mathbf{z}_{it}'\boldsymbol{\gamma}$ 用来描述无效率项对数中可以观测到的异质性, 未能观测到的异质性服从 $N(0, h_{it})$. 方差 h_{it} 为GARCH(p, q)模型, 当 $\lambda_1 = \dots = \lambda_p = \alpha_1 = \dots = \alpha_q = 0$ 时, h_{it} 恒为常数 λ_0 , 即无效率项 u_{it} 不存在异方差, 否则方差 h_{it} 随时间变化, 依赖于其以前时刻的方差和式(1)中的扰动项平方. 现有文献中较多的涉及无效率项中已观测到的异质性, 而对未观察到的异质性研究较少, 文[8]和[11]中假设未观测到的异质性服从一定分布. 本文在已有文献的基础上, 用协变量表示无效率项中可以观测到的异质性. 对未观察到的异质性, 通过假设对数无效率项方差为以前时刻方差和扰动平方项的线性组合, 讨论无效率项中未观测到的异质性及其变化规律.

记 $\boldsymbol{\lambda} = [\lambda_0, \lambda_1, \dots, \lambda_p]$, $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_q]$, $\mathbf{u} = [u_{it}, i = 1, 2, \dots, N, t = 1, \dots, T]$, 模型的待估参数为 $\boldsymbol{\theta} = [\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\lambda}, \sigma^2]$ 和隐含变量 $\mathbf{u}$. 设参数 $\boldsymbol{\theta}$ 的先验密度为 $p(\boldsymbol{\theta})$, 由Bayesian公式, 可以将参数估计模型看成是由 $p(y|\mathbf{x}, \mathbf{u}, \boldsymbol{\beta}, \sigma^2)$, $p(\mathbf{u}|\mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\lambda})$ 和 $p(\boldsymbol{\theta})$ 三部分组成的分层模型, 其中

$$\mathbf{x} = [\mathbf{x}_{it}, i = 1, 2, \dots, N, t = 1, \dots, T], \quad \mathbf{y} = [y_{it}, i = 1, 2, \dots, N, t = 1, \dots, T].$$

y 和 $\mathbf{u}$ 的联合分布为

$$p(y, \mathbf{u}|\mathbf{x}, \mathbf{z}, \boldsymbol{\theta}) = p(y|\mathbf{x}, \mathbf{u}, \boldsymbol{\beta}, \sigma^2)p(\mathbf{u}|\mathbf{z}, \boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\lambda})$$

$$\begin{aligned}
&= (2\pi\sigma^2)^{-(NT/2)} \exp \left[-\frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{t=1}^T (y_{it} + u_{it} - \mathbf{x}'_{it}\boldsymbol{\beta})^2 \right] \\
&\quad \times \prod_{i=1}^N \prod_{t=1}^T (h_{it})^{-1/2} \exp \left[-\frac{1}{2h_{it}} (\log u_{it} - \mathbf{z}'_{it}\boldsymbol{\gamma})^2 - \log u_{it} \right].
\end{aligned}$$

§3 Bayesian分析及抽样方法

本节对动态异方差随机前沿模型进行Bayesian分析并采用Gibbs抽样方法对模拟参数的后验均值进行推断. 首先确定参数的先验分布, 先验分布的选取应满足相容性原则, 即随着样本量的增大, 先验信息的影响减弱. 假设参数 $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\alpha}$, $\boldsymbol{\lambda}$ 和尺度参数 σ^2 相互独立, 则参数 $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\alpha}$, $\boldsymbol{\lambda}$ 的联合先验具有形式

$$p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\lambda}) = p(\boldsymbol{\beta}) p(\boldsymbol{\gamma}) p(\boldsymbol{\lambda}, \boldsymbol{\alpha}).$$

由于参数的先验信息很难得到, 而不恰当的先验会对模拟结果产生影响. 因此, 参数 $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\alpha}$, $\boldsymbol{\lambda}$ 的先验皆选为无信息先验. 另外, 为确保方差 h_{it} 为正和模型的平稳性, 系数 $\boldsymbol{\lambda}$ 和 $\boldsymbol{\alpha}$ 必须满足 $\lambda_i \geq 0$, $\alpha_i \geq 0$ 和

$$\sum_{i=1}^p \lambda_i + \sum_{i=1}^q \alpha_i < 1.$$

那么, 参数 $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\alpha}$, $\boldsymbol{\lambda}$ 的联合先验可以写为

$$p(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\lambda}) \propto \mathbf{I}\left(\sum_{i=1}^p \lambda_i + \sum_{i=1}^q \alpha_i < 1\right) \prod_{i=0}^p \mathbf{I}(\lambda_i \geq 0) \prod_{i=1}^q \mathbf{I}(\alpha_i \geq 0).$$

其中, $\mathbf{I}(\cdot)$ 为示性函数.

尺度参数 σ^2 的先验分布可取倒Gamma分布(参见[13]), 即

$$p(\sigma^2) \propto \sigma^{-p_0-1} \exp(-q_0/\sigma^2) \sim RG((p_0-1)/2, q_0).$$

其中, $p_0 > 0$ 和 $q_0 > 0$ 为先验分布的参数.

由Bayesian理论, 将参数的先验分布和样本信息综合可得到参数后验分布. $\boldsymbol{\theta}$ 和 $\mathbf{u}$ 的联合后验分布为

$$\begin{aligned}
p(\boldsymbol{\theta}, \mathbf{u} | \mathbf{x}, \mathbf{z}, y) &\propto p(y, \mathbf{u} | \mathbf{x}, \mathbf{z}, \boldsymbol{\theta}) p(\boldsymbol{\theta}), \\
&\propto \sigma^{-(NT+p_0+1)} \prod_{i=1}^N \prod_{t=1}^T (h_{it})^{-1/2} \exp \left[-\frac{1}{2\sigma^2} \left(\sum_{i=1}^N \sum_{t=1}^T (y_{it} + u_{it} - \mathbf{x}'_{it}\boldsymbol{\beta})^2 + 2q_0 \right) \right. \\
&\quad \left. + \sum_{i=1}^N \sum_{t=1}^T \left(-\frac{1}{2h_{it}} (\log u_{it} - \mathbf{z}'_{it}\boldsymbol{\gamma})^2 - \log u_{it} \right) \right] \mathbf{I}\left(\sum_{i=1}^p \lambda_i + \sum_{i=1}^q \alpha_i < 1\right) \\
&\quad \times \prod_{i=0}^p \mathbf{I}(\lambda_i \geq 0) \prod_{i=1}^q \mathbf{I}(\alpha_i \geq 0). \tag{2}
\end{aligned}$$

为对模型进行预测或参数推断, Bayesian模型往往需要对高维概率密度分布进行积分. MCMC是一种解决高维积分问题的迭代Monte Carlo方法, 它解决了复杂表达式难以进行高维积分的问题, 应用非常广泛. MCMC通过利用Markov链进行积分得到一系列的Markov链模拟值, 该链的平稳分布收敛到已知参数的后验分布, 即可看作后验分布的独立样本. Gibbs抽

样是一种最常见且最简单的MCMC方法, 这里使用Gibbs抽样方法对参数后验均值进行推断. 为了使MCMC方法产生的Markov链的平稳分布是(2), 下面讨论Gibbs抽样方法的具体策略. 设Gibbs抽样得到的Markov链

$$\{\beta^{(i)}, \sigma^{2(i)}, \gamma^{(i)}, \lambda^{(i)}, \alpha^{(i)}, \mathbf{u}^{(i)}, i = 1, \dots, M\}$$

收敛到已知的参数后验分布. 因此, 参数后验期望可以通过Gibbs抽样得到的平均值来近似, 即

$$E[\hat{f}(\theta)] = \frac{1}{M-m} \sum_{i=m+1}^M f(\theta^{(i)}).$$

Gibbs抽样的具体迭代步骤如下: 给定迭代起始点 $\{\beta^{(1)}, \sigma^{2(1)}, \gamma^{(1)}, \lambda^{(1)}, \alpha^{(1)}, \mathbf{u}^{(1)}\}$, 依次取 $i = 2, \dots, M$, 那么第 i 次迭代的实现步骤如下:

步1 从分布 $p(\beta^{(t)} | \sigma^{2(t-1)}, \gamma^{(t-1)}, \lambda^{(t-1)}, \alpha^{(t-1)}, \mathbf{u}^{(t-1)}, y, \mathbf{x})$ 里产生 $\beta^{(t)}$,

步2 从分布 $p(\sigma^{2(t)} | \beta^{(t)}, \gamma^{(t-1)}, \lambda^{(t-1)}, \alpha^{(t-1)}, \mathbf{u}^{(t-1)}, y, \mathbf{x})$ 里产生 $\sigma^{2(t)}$,

步3 从分布 $p(\mathbf{u}^{(t)} | \beta^{(t)}, \sigma^{2(t)}, \gamma^{(t-1)}, \lambda^{(t-1)}, \alpha^{(t-1)}, y, \mathbf{x})$ 里产生 $\mathbf{u}^{(t)}$,

步4 从分布 $p(\gamma^{(t)} | \beta^{(t)}, \sigma^{2(t)}, \lambda^{(t-1)}, \alpha^{(t-1)}, \mathbf{u}^{(t)}, y, \mathbf{x})$ 里产生 $\gamma^{(t)}$

步5 从分布 $p(\lambda^{(t)} | \beta^{(t)}, \sigma^{2(t)}, \gamma^{(t)}, \alpha^{(t-1)}, \mathbf{u}^{(t)}, y, \mathbf{x})$ 里产生 $\lambda^{(t)}$,

步6 从分布 $p(\alpha^{(t)} | \beta^{(t)}, \sigma^{2(t)}, \gamma^{(t)}, \lambda^{(t)}, \mathbf{u}^{(t)}, y, \mathbf{x})$ 里产生 $\alpha^{(t)}$,

步7 对 $h = [h_{it}, i = 1, 2, \dots, N, t = 1, \dots, T]$ 进行更新.

下面对参数抽样的具体策略进行分析. 由(2)式可得到每个参数后验分布的核.

$$\begin{aligned} p(\beta | \mathbf{u}, \gamma, \lambda, \alpha, y, \mathbf{x}) &\propto \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{t=1}^T (y_{it} + u_{it} - \mathbf{x}'_{it}\beta)^2\right\} \\ &\propto N((\mathbf{x}'\mathbf{x})^{-1}\mathbf{x}'(y + \mathbf{u}), \sigma^2(\mathbf{x}'\mathbf{x})^{-1}), \\ p(\sigma^2 | y, \mathbf{x}, \mathbf{u}, \beta, \gamma, \lambda, \alpha) &\propto \sigma^{-(NT+p_0+1)} \exp\left\{-\frac{1}{2\sigma^2} \left[\sum_{i=1}^N \sum_{t=1}^T (y_{it} + u_{it} - \mathbf{x}'_{it}\beta)^2 + 2q_0\right]\right\} \\ &\propto \left(\frac{(y + \mathbf{u} - \mathbf{x}'\beta)'(y + \mathbf{u} - \mathbf{x}'\beta) + 2q_0}{\sigma^2}\right)^{\frac{NT+p_0+3}{2}-1} \\ &\quad \times \exp\left\{-\frac{(y + \mathbf{u} - \mathbf{x}'\beta)'(y + \mathbf{u} - \mathbf{x}'\beta) + 2q_0}{2\sigma^2}\right\}. \end{aligned} \quad (3)$$

其中, $\mathbf{x}$ 为 $(N \cdot T) \times k$ 矩阵, y 为 $(N \cdot T) \times 1$ 矩阵. 由(3)式可知,

$$((y + \mathbf{u} - \mathbf{x}'\beta)'(y + \mathbf{u} - \mathbf{x}'\beta) + 2q_0)/\sigma^2 \sim Ga((NT + p_0 + 3)/2, 1/2) \doteq \chi^2_{NT+p_0+3}.$$

其中, $\chi^2_{NT+p_0+3}$ 表示自由度为 $NT + p_0 + 3$ 的 χ^2 分布.

$$p(\gamma | \mathbf{u}, \mathbf{z}, \lambda, \alpha) \propto \exp\left\{-\sum_{i=1}^N \sum_{t=1}^T \frac{(\log u_{it} - \mathbf{z}'_{it}\gamma)^2}{2h_{it}}\right\} \propto N(\mu_\gamma, \Sigma_\gamma).$$

可知 γ 的后验分布是正态分布, 其中

$$\mu_\gamma = \Sigma_\gamma \left(\sum_{i=1}^N \sum_{t=2}^T \mathbf{z}_{it} h_{it}^{-1} \log u_{it}\right), \Sigma_\gamma = \left(\sum_{i=1}^N \sum_{t=1}^T \mathbf{z}_{it} h_{it}^{-1} \mathbf{z}'_{it}\right)^{-1}.$$

参数 β, γ 和 σ^2 的后验条件分布都为标准分布, 可直接进行抽样.

隐含变量 u_{it} 的后验分布随着时间变化, 第 i 个公司 t 时刻无效率项 u_{it} 的后验分布为

$$p(u_{it}|y_{it}, \mathbf{x}_{it}, \mathbf{z}_{it}, \boldsymbol{\beta}, \boldsymbol{\gamma}, h) \propto \frac{1}{u_{it}} \exp\left\{-\frac{1}{2\sigma^2}(y_{it} + u_{it} - \mathbf{x}_{it}'\boldsymbol{\beta})^2 - \frac{1}{2h_{it}}(\log u_{it} - \mathbf{z}_{it}'\boldsymbol{\gamma})^2\right\}.$$

其后验分布不是标准分布, 不能进行直接抽样, 可以通过切片抽样或舍选抽样得到. 文中使用的是切片抽样方法, 产生 u_{it} 的具体步骤见算法1.

算法1(产生 u_{it})

步1 从 $U[0, f_1(u_{it}^{(t-1)})]$ 中抽取 ω_1 .

步2 从 $U[0, f_2(u_{it}^{(t-1)})]$ 中抽取 ω_2 .

步3 从 U_A 中抽取 $u_{it}^{(t)}$. 其中 U_A 为 A 上的均匀分布, $A = \{x; f_i(x) \geq \omega_i, i = 1, 2\}$.

算法1中 f_1 和 f_2 是 u_{it} 后验密度函数的分解, 满足

$$p(u_{it}|y_{it}, \mathbf{x}_{it}, \mathbf{z}_{it}, \boldsymbol{\beta}, \boldsymbol{\gamma}, h) = f_1(u_{it})f_2(u_{it}).$$

文中令

$$f_1(u_{it}) = \exp\{-(y_{it} + u_{it} - \mathbf{x}_{it}'\boldsymbol{\beta})^2/(2\sigma^2)\}, f_2(u_{it}) = \exp\left\{-\frac{1}{2h_{it}}(\log u_{it} - \mathbf{z}_{it}'\boldsymbol{\gamma})^2 - \log(u_{it})\right\}.$$

参数 $\boldsymbol{\lambda}$ 和 $\boldsymbol{\alpha}$ 的后验分布为

$$p(\boldsymbol{\lambda}, \boldsymbol{\alpha}|\mathbf{u}, \boldsymbol{\gamma}) \propto \prod_{i=1}^N \prod_{t=1}^T (h_{it})^{-\frac{1}{2}} \exp\left\{-\sum_{i=1}^N \sum_{t=1}^T \frac{(\log u_{it} - \mathbf{z}_{it}'\boldsymbol{\gamma})^2}{2h_{it}}\right\}.$$

其中

$$h_{it} = \lambda_0 + \sum_{j=1}^p \lambda_j \varepsilon_{i,t-j}^2 + \sum_{j=1}^q \alpha_j h_{i,t-j}.$$

从后验分布可知, 在给定参数 $\mathbf{u}, \boldsymbol{\gamma}$ 的前提下, 参数 $\boldsymbol{\lambda}$ 和 $\boldsymbol{\alpha}$ 中给定任意一个参数的值, 另一个参数的条件后验核密度函数都不是标准分布密度函数. 在这里取样方法参照Nakatsuma(参见[12])所提议的近似GARCH模型, 用Metropolis-Hastings抽样方法对其进行抽样. 令

$$w_{it} = \varepsilon_{it}^2 - h_{it},$$

式

$$h_{it} = \lambda_0 + \sum_{j=1}^p \lambda_j \varepsilon_{i,t-j}^2 + \sum_{j=1}^q \alpha_j h_{i,t-j}$$

两端同时减去 ε_{it}^2 , 整理得近似GARCH模型:

$$\varepsilon_{it}^2 = \lambda_0 + \sum_{j=1}^l (\lambda_j + \alpha_j) \varepsilon_{i,t-j}^2 + w_{it} - \sum_{j=1}^q \alpha_j w_{i,t-j}, \quad w_{it} \sim N(0, 2h_{it}^2).$$

其中 $l = \max\{p, q\}$, 且有

$$E(w_{it}|\varepsilon_{it}^2, h_{it}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) = 0, \text{Var}(w_{it}|\varepsilon_{it}^2, h_{it}, \boldsymbol{\lambda}, \boldsymbol{\alpha}) = 2h_{it}^2.$$

以 $\boldsymbol{\lambda}$ 的取样为例, Metropolis-Hastings抽样方法的具体步骤为:

算法2(产生 $\boldsymbol{\lambda}$)

步1 从建议分布 $\pi(\boldsymbol{\lambda})$ 中抽取 $\boldsymbol{\lambda}'$.

步2 从 $U(0, 1)$ 中抽取 e .

步3 若 $e \leq \alpha(\lambda^{(t-1)}, \lambda')$, 则 $\lambda^{(t)} = \lambda'$; 否则, $\lambda^{(t)} = \lambda^{(t-1)}$; 其中

$$\alpha(\lambda^{(t-1)}, \lambda') = \min \left\{ 1, \frac{p(\lambda'|\beta, \sigma^2, \alpha, u, y, x)\pi(\lambda^{(t-1)})}{p(\lambda^{(t-1)}|\beta, \sigma^2, \alpha, u, y, x)\pi(\lambda')} \right\}.$$

算法2中 λ 的建议分布需要主观设定, 近似GARCH的似然函数可以写为:

$$\begin{aligned} \pi(\varepsilon^2|\lambda, \alpha, u, \gamma) &\propto \prod_{i=1}^N \prod_{t=1}^T \frac{1}{h_{it}} \exp\left\{-\frac{w_{it}^2}{4h_{it}^2}\right\} \\ &\propto \prod_{i=1}^N \prod_{t=1}^T \frac{1}{h_{it}} \exp\left\{-\frac{(\varepsilon_{it}^2 - \zeta_{it}\lambda')^2}{4h_{it}^2}\right\} \\ &\propto N(\hat{\mu}_\lambda, \hat{\Sigma}_\lambda). \end{aligned}$$

其中,

$$\zeta_{it} = [\tilde{l}_{it}, \tilde{\varepsilon}_{i,t-1}^2, \dots, \tilde{\varepsilon}_{i,t-p}^2], \tilde{l}_{it} = 1 + \sum_{j=1}^q \alpha_j \tilde{l}_{i,t-j}, \tilde{\varepsilon}_{it}^2 = \varepsilon_{it}^2 + \sum_{j=1}^q \alpha_j \tilde{\varepsilon}_{i,t-j}^2.$$

那么, λ 建议分布为

$$\lambda|\alpha, u, \gamma \sim N(\hat{\mu}_\lambda, \hat{\Sigma}_\lambda) \prod_{i=0}^p I(\lambda_i \geq 0).$$

这里,

$$\hat{\mu}_\lambda = \hat{\Sigma}_\lambda \sum_{i=1}^N \sum_{t=1}^T \frac{\varepsilon_{it}^2 \zeta'_{it}}{2h_{it}^2}, \hat{\Sigma}_\lambda = \left(\sum_{i=1}^N \sum_{t=1}^T \frac{\zeta'_{it} \zeta_{it}}{2h_{it}^2} \right)^{-1}.$$

为生成 α 的建议分布, 将 $w_{it}(\alpha)$ 在 α^* 处用Taylor展式展至一次项, 则近似GARCH模型的似然函数为

$$\begin{aligned} \pi(\alpha|\lambda, u, \gamma) &\propto \prod_{i=1}^N \prod_{t=1}^T \frac{1}{h_{it}} \exp\left\{-\frac{w_{it}^2}{4h_{it}^2}\right\} \\ &\propto \prod_{i=1}^N \prod_{t=1}^T \frac{1}{h_{it}} \exp\left\{-\frac{(w_{it}(\alpha^*) - \zeta_{it}(\alpha - \alpha^*))^2}{4h_{it}^2}\right\} \\ &\propto N(\hat{\mu}_\alpha, \hat{\Sigma}_\alpha). \end{aligned}$$

上式中 α^* 是 α 的加权非线性最小二乘估计值, 即

$$\alpha^* = \arg \min \sum_{i=1}^N \sum_{t=1}^T \frac{w_{it}(\alpha)}{2h_{it}^2}, (0 < \alpha_j < 1, j = 1, \dots, q).$$

$\zeta_{it} = [\zeta_{it1}, \dots, \zeta_{itq}]$, 其中 ζ_{itj} 是 w_{it} 在点 α_j^* 处的一阶导数值的相反数. 又

$$\frac{dw_{it}(\alpha_j)}{d\alpha_j} = -h_{i,t-j} + \sum_{s=1}^q \alpha_s \frac{dw_{i,t-s}}{d\alpha_j},$$

则有

$$\zeta_{itj} = -\frac{dw_{it}(\alpha_j)}{d\alpha_j} = h_{i,t-1}(\alpha^*) + \sum_{s=1}^q \alpha_s^* \zeta_{i,t-s,j}.$$

那么, 参数 α 的建议分布为

$$\alpha | \lambda, \gamma, \mathbf{u} \sim N(\hat{\mu}_\alpha, \hat{\Sigma}_\alpha) \prod_{i=1}^q I(\alpha_i \geq 0).$$

令

$$Y_{it}^\alpha = w_{it}(\alpha^*) + \zeta_{it}(\alpha^*)', X_{it}^\alpha = \zeta_{it},$$

则

$$\hat{\mu}_\alpha = \hat{\Sigma}_\alpha \sum_{i=1}^N \sum_{t=2}^T \frac{Y_{it}^\alpha X_{it}^\alpha}{2h_{it}^2}, \hat{\Sigma}_\alpha = \left(\sum_{i=1}^N \sum_{t=2}^T \frac{(X_{it}^\alpha)'(X_{it}^\alpha)}{2h_{it}^2} \right)^{-1}.$$

§4 模拟实验

为验证Gibbs抽样方法的效果, 这里进行了模拟实验. 为方便计算, 令 $p = q = 1$ 并对两个公司的情况进行模拟, 考虑如下模型:

$$y_{it} = \mathbf{x}_{it}'\beta + v_{it} - u_{it}, \quad i = 1, 2, \quad t = 1, \dots, T,$$

$$\log u_{it} = \mathbf{z}_{it}'\gamma + \varepsilon_{it}, \quad \varepsilon_{it} = \sqrt{h_{it}}\xi_{it}, \quad h_{it} = \lambda_0 + \lambda_1 \varepsilon_{i,t-1}^2 + \alpha_1 h_{i,t-1}.$$

取 $k = 3$, $m = 3$, $\mathbf{x}_{it}$, $\mathbf{z}_{it}$ 的第一列均取1, 为表现出相关性, $\mathbf{x}_{it}$, $\mathbf{z}_{it}$ 的剩余 $k - 1$ 列的元素从二元标准正态分布中随机产生, 两者的相关系数为0.5, 如文[14]. 取参数 β 真值为 $[1, 1, 1]'$, γ 真值为 $[-0.3, 1, 1]'$. 取 $\sigma^2 = 0.05$, $\lambda_0 = 0.05$, $\lambda_1 = 0.2$, $\alpha_1 = 0.7$. 对于先验分布的参数分别取 $p_0 = 4$, $q_0 = 0.01$. 分别取样本量100, 500和1000三种情况来进行模拟比较, 模拟轨道长度60000并抛去轨道的前10000个点. 各参数后验估计值和后验标准差见表1.

		表1 参数后验均值结果					
		T=100		T=500		T=1000	
参数	真值	均值	sd	均值	sd	均值	sd
β_1	1	1.083	0.067	1.015	0.025	1.040	0.019
β_2	1	0.977	0.026	0.972	0.012	1.010	0.009
β_3	1	1.053	0.031	1.004	0.012	1.012	0.009
γ_1	-0.3	-0.184	0.135	-0.282	0.051	-0.239	0.038
γ_2	1	0.917	0.086	0.949	0.034	0.996	0.026
γ_3	1	0.949	0.101	1.003	0.033	0.984	0.026
σ^2	0.05	0.036	0.011	0.037	0.004	0.047	0.004
λ_0	0.05	0.133	0.066	0.054	0.026	0.048	0.011
λ_1	0.2	0.209	0.096	0.124	0.041	0.120	0.021
α_1	0.7	0.584	0.131	0.727	0.096	0.774	0.035

由表1, 对比各个参数的真值和估计值, 参数 β 的估计值在样本量很小时就非常接近真实值, 参数 $\mathbf{z}$ 的估计值样本量很小时就很接近真值. 参数 λ_0 , λ_1 , α_1 和 σ^2 在样本量小时接近真值, 参数的估计随着样本量的增大, 模拟效果变好. 参数的后验标准差随着样本量的增大明显减小, 特别是参数 λ_0 , λ_1 和 α_1 , 在样本量较小时标准差较大, 随着样本量的增大明显减小. 可见, 用Gibbs抽样对动态异方差随机前沿模型进行Bayesian分析有比较好的效果. 在模拟实验过程中发现, 当样本量较小特别是公司多时间短的情况下, λ_0 的模拟值会出现均非常接近于零的情况, 使 h_{it} 的后验均值几乎为零, 即参数估计结果中无效率项方差估计值非常小. 这是因为, λ_0 的抽样采用的是Metropolis-Hasting算法. 在抽样过程中, 一旦在一步中出现 λ_0 的模拟值很接近于零时, λ_0 的建议分布的方差将很小, 使得Metropolis-Hasting抽样更新效率极低. 而当样本抽样时间区间很

短时, 极易出现在某一步迭代中 λ_0 的模拟值非常接近于零的情况. 将 λ_0 的先验由无信息先验变为均匀分布 $U[\delta, 1]$, 其中 δ 为一很小的正数, 可避免 λ_0 的模拟值几乎为零的情况.

§5 实证研究

将本文模型应用于哥伦比亚电力改革. 样本数据来自文[15]中, 取自1998-2012年哥伦比亚的电力分销公司, 本文对原始数据进行筛选, 取出其中时间序列较长的12 家公司共174个观察值. 考虑如下随机前沿模型,

$$\begin{aligned} -\log y_{it} = & \beta_0 + \beta_2 \log x_{1it} + \beta_3 \log x_{2it} + \frac{1}{2}\beta_4 \log x_{1it} \log x_{1it} + \frac{1}{2}\beta_5 \log x_{2it} \log x_{2it} \\ & + \beta_6 \log x_{1it} \log x_{2it} + v_{it} - u_{it}, \\ \log u_{it} = & \gamma_0 + \gamma_1 z_{it} + \varepsilon_{it}. \end{aligned}$$

其中, y_{it} , x_{1it} , x_{2it} 分别表示第*i*家公司*t*时刻的输入总费用, 输出总能量和消费者数量. 无效率项异方差变量 z_{it} 为第*i*家公司*t*时刻的用户密度. 数据和模型的进一步解释可参见文[16]. 先验参数取值与模拟实验部分的值一致, p 和 q 的值均取1. Gibbs抽样模拟轨道长60000并抛去模拟轨道的前10000个点. 主要参数的后验均值和标准差见表2, 其中括号中数字为参数后验标准差.

表2 参数后验均值结果

β_0	β_1	β_2	β_3	β_4	β_5
-18.403	-0.574	-0.871	0.869	1.281	-1.028
(0.433)	(0.292)	(0.244)	(0.124)	(0.195)	(0.153)
γ_0	γ_1	σ^2	λ_0	λ_1	α_1
-6.058	0.102	0.017	0.012	0.192	0.160
(0.893)	(0.209)	(0.001)	(0.003)	(0.136)	(0.148)

由表2可知, 从 γ 的估计值可以看出, 公司的用户密度会对效率产生一定的影响. 参数 λ_0 的估计值为0.011, 而 λ_1 和 α_1 的估计值分别为0.192 和0.160, 说明对数无效率项的方差依赖于过去误差项和滞后条件误差. 可见, 随机前沿模型中不能忽略公司特征引起的异方差和方差的时变性. 实例主要研究了条件方差 h_{it} 为GARCH(1,1)的情况, 未对高阶的情况进行深入研究. 另外, 可以将本文与文[8]和[11]进行结合, 对无效率项进行更全面的诠释.

参考文献:

- [1] Aigner D, Knox-Lovell C A, Schmidt P. Formulation and estimation of stochastic frontier production function models [J]. J Econometrics, 1977, 6: 21-37.
- [2] Meeusen W, van der Broeck J. Efficiency estimation from Cobb-Douglas production functions with Composed Error[J]. Internat Econom Rev, 1977, 18: 435-444.
- [3] Stevenson R E. Likelihood functions for generalized stochastic frontier model[J]. J Econometrics, 1980, 13: 57-66.
- [4] Greene W H. A gamma-distributed stochastic frontier model[J]. J Econometrics, 1994, 61: 273-303.
- [5] 刘晓君, 张世斌. 正态倒Gamma随机前沿模型的Bayesian推断[J]. 高校应用数学报, 2013, 28(4): 488-496.
- [6] Greene W. Reconsidering heterogeneity in panel data estimators of the stochastic frontier model[J]. J Econometrics, 2005, 126: 269-303.

- [7] Carta A, Steel M F J. Modelling multi-output stochastic frontiers using copulas[J]. Comput Statist Data Anal, 2012, 56: 3757-3773.
- [8] Tsionas E. Inference in dynamic stochastic frontier models[J]. J Appl Econom, 2006, 21: 669-676.
- [9] Chen Yi-Yi, Schmidt P, Wang Hung-Jen. Consistent estimation of the fixed effects stochastic frontier model[J]. J Econometrics, 2014, 181: 65-76.
- [10] Griffin J, Steel M. Flexible mixture modelling of stochastic frontiers[J]. J Prod Anal, 2008, 29: 33-50.
- [11] Galón J, Veiga H, Wiper M. Bayesian estimation of inefficiency heterogeneity in stochastic frontier models[J]. J Prod Anal, 2014, 42: 85-101.
- [12] Nakatsuma T. Bayesian Analysis of ARMA-GARCH Models: A Markov chain Sampling Approach[J]. J Econometrics, 2000, 95: 57-69.
- [13] Fernandez C, Osiewalski J, Steel M F J. On the use of panel data in stochastic frontier models[J]. J Econometrics, 1997, 79: 169-193.
- [14] Wang Hung-Jen, Schmidt P. One step and two step estimation of the effects of exogenous variables on technical efficiency levels[J]. Prod Anal, 2002, 18: 129-144.
- [15] Galón J E, Pollitt M. Inefficiency persistence and heterogeneity in Colombian electricity utilities[J]. Energ Econ, 2014, 46: 31-44.
- [16] Growitsch T, Jamasb T, Pollitt M. Quality of service, efficiency and scale in network industries: an analysis of European electricity distribution[J]. Appl Econ, 2005, 41: 2555-2570.

Bayesian inference for dynamic heterogeneity stochastic frontier model

CHENG Di¹, ZHANG Shi-bin²

(1. School of Math. Sci., Inner Mongolian Univ., Hohhot 010021, China;

2. Dept. of Math., Shanghai Maritime Univ., Shanghai 201306, China)

Abstract: If heterogeneity of the “inefficiency” term is disregarded, it will result in the incorrect estimate of this term in the stochastic frontier model. By combining the influence from characteristic differences of individuals with the time-varying property of variance, a dynamic heterogeneity stochastic frontier model is proposed. By the Gibbs sampling, the methodology for Bayesian analysis of the dynamic heterogeneity stochastic frontier model is given. For each model parameter, the posterior distribution is derived. A simulation study shows that under the criterion of minimizing the posterior mean square error, the Bayesian estimate is close to its true value for small and medium sized samples. From the Bayesian analysis based on the real electric power company generation data, it is evidenced that there exists the time-varying property for the variance of the logarithm “inefficiency” term.

Keywords: stochastic frontier model; Bayesian inference; heterogeneity; Gibbs sampling; Metropolis-Hastings sampling

MR Subject Classification: 62F15; 65C60