arXiv:2301.07245v1 [math.ST] 18 Jan 2023

稳健且高效的 Breusch-Pagan 检验统计量：将 β-分数拉格朗日乘子检验应用于非同分布个体

Nirian Martín，马德里康普顿斯大学（nimartin@ucm.es）

2023年1月19日

摘要

在计量经济学中，Breusch-Pagan 检验统计量已成为拉格朗日乘数（LM）检验的标志性应用。本文将介绍用于线性回归模型异方差检验的 β-分数 LM 检验，该检验通过调节参数 $\beta \geq 0$ 在稳健性与效率之间进行权衡；当不存在离群值时，$\beta = 0$ 经典 Breusch-Pagan检验统计量效率最高。该检验具有非常简洁优雅的表达式，并可作出引人入胜的最小二乘法解释。对于复合原假设，我们将 Basu 等人（2022）的方法从同分布个体推广至非同分布个体，以构造该检验统计量。文中在正态性假设下，详细论证了其稳健性和效率性质。我们还推导出一个修正版，即 Koenker 的 β-分数检验统计量。本文通过一个经典示例，说明在实际中如何处理线性回归模型的新型异方差检验。

JEL 分类：C01，C10；
关键词：布鲁施-帕甘检验、复合原假设、密度幂散度、异方差、拉格朗日乘子检验、线性回归、稳健性。

1 引言

大多数本科计量经济学教材都将 Breusch-Pagan 拉格朗日乘数（LM）检验列为线性回归模型中最重要的异方差检验之一。该检验由 Breusch 和 Pagan（1979）提出，Godfrey 也独立推导出了这一检验。

Godfrey（1978a）是在乘法异方差这一更强假设下推导出该检验的。此后，Koenker（1981）以及 Koenker 和 Basset（1982）提出了该检验的修订版本，放宽了误差分布正态性的假设，从而适用于更一般的分布。针对特定类型的异方差，若设计矩阵具有特定结构，并加入回归变量的交叉乘积项，White（1980）据此提出了一种异方差检验；但 Waldman（1983）指出，根据对误差项采用正态或非正态分布假设，White 版本的检验是 Breusch-Pagan 检验或 Koenker 检验的一个特例。上述检验在面对异常值数据时稳健性不足。例如，Carrol 和 Ruppert（1988，第98页）、Green（第314页）、Kmenta（1986，第295页）、Lyon 和 Tsay（1996，第339页）以及 Kalina（2011）都报告了这一问题。近期，Berenguer-Rico 和 Wilms（2021）建议在使用 White 检验统计量检验异质性时剔除离群值；而 Alih 和 Ong（2015）则提出了一种稳健版 Goldfeld-Quandt 检验统计量，以替换其中非稳健的部分。

许多统计程序的稳健性往往是以牺牲效率为代价实现的。兼具渐近效率和稳健性的估计量与检验统计量颇具吸引力。这一思想最早由 Beran（1977）通过海灵格距离引入估计问题。Kolmogorov（1933）和 Mahalanobis（1936）分别以不同方式提出了基于样本构造的两个概率分布之间距离（或散度）的原始概念。在参数统计推断中，从一开始到后续发展，提出的新统计距离都着眼于经验分布 g与基于模型的分布 $f_{\boldsymbol{\theta}}$之间的接近程度，其中 $\boldsymbol{\theta}$ 是我们关注的 p维参数向量。当真实分布属于基于模型的分布时，众所周知，Kullback-Leibler（1951）散度涵盖了与最大似然估计（MLE）及其相关检验统计量有关的全部经典统计理论。Csisz´ar（1967）提出了一大类ϕ-散度度量，由此可得到 BAN（最佳渐近正态）估计量，即渐近而言与最大似然估计量同样有效率的估计量。上述距离中的一个特例是 Cressie 和 Read（1984）所称的幂散度，而后者又将海灵格距离作为一个特例。更多详情见 Pardo（2006）和 Basu 等人（2011）。随着文献不断发展，人们提出了若干新的稳健且有效率的最小距离程序，其中 Basu 等人（1998）提出的密度幂散度（DPD）度量已广受欢迎。相较于海灵格距离，这些方法的新颖之处在于估计量具有有界影响函数，并且其适用范围扩展至连续总体，而此前的研究大多只处理离散总体。此后，DPD 在统计推断中日益受到关注，并被应用于不同参数模型中的估计与检验。对于线性回归，从条件角度或固定设计矩阵的角度出发，Ghosh 和 Basu（2013）首次正式阐明了如何在样本观测值相互独立但并非同分布的情况下，将 DPD 用于估计。

在该样本中。对于线性回归模型中基于 DPD 的稳健检验，Ghosh等人（2016）的第6.3节提出了 β-Wald 检验统计量，而 Qin和Priebe（2017）则引入了 β-似然比检验统计量（Lq-似然比型检验）。与 β-Wald 检验统计量相比， β-似然比检验统计量存在一个明显缺点：其渐近分布取决于一些权重，这些权重需要计算依赖于未知参数的矩阵的特征值；除标量参数的情况外，这会影响检验统计量计算的准确性。根据 Basu 等人（2022），本文将要发展的 β-分数 LM 检验预计具有与 β-Wald 检验统计量相同的渐近分布。此外，正如 Boos（1992）指出的，广义得分检验（包括 β-分数 LM 检验）的优势在于其不变性。这正是 Breusch-Pagan LM 检验的特点，因为它适用于检验一般类型异方差的存在。一个统一的检验统计量表达式涵盖了异方差函数的不同可能形式，而其他类型的检验统计量则无法做到这一点。

在计量经济学中，已有大量研究基于最小 DPD 估计量，在保留较高效率的同时评估稳健性（例如，参见Lee和Song（2009）、Kim和Lee（2013、2017、2018））。据我们所知，除拟合优度检验情形外（例如，Kim（2018）），尚无关于使用此类估计量进行检验的研究。本文基于最小 DPD 估计量，考察了相合估计量以及无偏估计方程。本文提出的方法在理论上具有双重稳健性，既体现在估计量的构造上，也体现在检验统计量上。该方法的主要特点一方面是通过有界的影响函数实现估计和检验的稳健性；随着调节参数 β增大，总体误差敏感度也随之增加，同时通过同一调节参数 β 趋近于 0，确保效率处于可接受水平。另一方面，本文还推导了Pitman渐近相对效率（渐近相对效率），它衡量在处理纯净数据时，为获得特定的渐近功效，按相对样本量计所需付出的代价。

本文结构安排如下。第2节介绍并阐述一种用于复合假设检验的新抽样框架，该框架基于相互独立但不同分布的观测值，并将其与异方差线性回归模型联系起来。第3节给出主要理论结果；第 β节推导 Breusch-Pagan4分数 LM 检验和影响函数分析；第5节推导其修正版，即 Koenker 的 β分数检验统计量。第7节通过一个著名示例，实际演示如何处理线性回归模型中的异质性新检验。第8节给出一些总结性评论。大部分证明汇集于本文末尾的附录中。

2 基本模型与检验设定

基于线性回归的条件版本，预测变量被视为固定，因此响应样本 $Y_i, i = 1, \ldots, n$ 需要对 Basu 等人（2002）针对非同分布个体提出的既有理论进行调整。在此情形下，设 $g_i(\cdot)$ 是 $i = 1, \ldots, n$ 的真实概率密度函数（p.d.f.），$f_{i,\boldsymbol{\theta}}(\cdot)$ 是模型在 $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$ 下的概率密度函数，则根据 Ghosh 和 Basu（2013），整个样本的密度幂散度为 $H_{n,\beta}(\boldsymbol{\theta}) = \sum_{i=1}^{n} d_\beta(g_i, f_{i,\boldsymbol{\theta}})$，其中

$$d_\beta(g_i, f_{i,\boldsymbol{\theta}}) = \int_{\mathcal{Y}} \left(f_{i,\boldsymbol{\theta}}^{\beta+1}(y) - \tfrac{\beta+1}{\beta} f_{i,\boldsymbol{\theta}}^{\beta}(y) g_i + \tfrac{1}{\beta} g_i^{\beta+1} \right) dy, \quad \beta > 0.$$

其中，$\mathcal{Y}$ 是所有响应变量样本的共同支撑集。实践中，未知的概率密度函数 $g_i(\cdot)$ 通过其经验形式 $\widehat{g}_i(\cdot)$ 进行近似。

H_0： $\boldsymbol{\theta} \in \Theta_0$ 与 H_1： $\boldsymbol{\theta} \in \Theta - \Theta_0$，其中

$$\Theta_0 = \{\boldsymbol{\theta} \in \Theta : \boldsymbol{m}(\boldsymbol{\theta}) = \mathbf{0}_r\}, \tag{1}$$

其中，$\boldsymbol{m}$： $\Theta \longrightarrow \mathbb{R}^r$，其中 $r < p$ 是复合原假设检验。假设定义 Θ_0 的函数满足一些正则条件，

$$\boldsymbol{M}(\boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \boldsymbol{m}^T(\boldsymbol{\theta}),$$

在 $\boldsymbol{\theta}$ 中存在且连续，并且 $\text{rank}(\boldsymbol{M}(\boldsymbol{\theta})) = r$。

定义 1 受限于（ $\boldsymbol{\theta}$所确立的原假设的最小 *DPD* 估计量*1*），可表示为

$$\widetilde{\boldsymbol{\theta}}_{n,\beta} = \arg\min_{\boldsymbol{\theta} \in \Theta_0} \widehat{H}_{n,\beta}(\boldsymbol{\theta}) = \arg\max_{\boldsymbol{\theta} \in \Theta_0} \left(\tfrac{\beta+1}{\beta} \tfrac{1}{n} \sum_{i=1}^{n} f_{i,\boldsymbol{\theta}}^{\beta}(Y_i) - \sum_{i=1}^{n} \mathrm{E}\left[f_{i,\boldsymbol{\theta}}^{\beta}(Y_i) \right] \right),$$

其中

$$\widehat{H}_{n,\beta}(\boldsymbol{\theta}) = \sum_{i=1}^{n} d_\beta(\widehat{g}_i, f_{i,\boldsymbol{\theta}}),$$

$$d_\beta(\widehat{g}_i, f_{i,\boldsymbol{\theta}}) = \mathrm{E}\left[f_{i,\boldsymbol{\theta}}^{\beta}(Y_i) \right] - \tfrac{\beta+1}{\beta} \tfrac{1}{n} f_{i,\boldsymbol{\theta}}^{\beta}(Y_i) + \tfrac{1}{\beta} \tfrac{1}{n}.$$

$\boldsymbol{\theta}$ 的受限最大似然估计量（最大似然估计量）是最小 DPD 估计量中的一个，因为

$$\widetilde{\boldsymbol{\theta}}_{\beta=0} = \arg\min_{\boldsymbol{\theta} \in \Theta_0} \lim_{\beta \to 0^+} \widehat{H}_{n,\beta}(\boldsymbol{\theta}) = \arg\max_{\boldsymbol{\theta} \in \Theta_0} \sum_{i=1}^{n} \log f_{i,\boldsymbol{\theta}}(Y_i) = \widetilde{\boldsymbol{\theta}}.$$

命题1 在（ $\boldsymbol{\theta}$）建立的原假设下受限的最小 *DPD* 估计量，通过求解以下包含 个方程和未知参数的方程组得到：*1), is obtained as solution in* $(\boldsymbol{\theta}^T, \boldsymbol{\lambda}_{n,\beta}^T)^T$ *of the following system of* $p+r$ *equations and unknown parameters*

$$\boldsymbol{U}_{n,\beta}(\boldsymbol{\theta}) - \boldsymbol{M}(\boldsymbol{\theta})\boldsymbol{\lambda}_{n,\beta} = \mathbf{0}_p, \tag{2}$$

$$\boldsymbol{m}(\boldsymbol{\theta}) = \mathbf{0}_r, \tag{3}$$

其中 $\boldsymbol{\lambda}_{n,\beta}$ 是一个 r 维拉格朗日乘子向量，其最小 *DPD* 估计量记为 $\widetilde{\boldsymbol{\lambda}}_{n,\beta}(\boldsymbol{\theta}_\beta)$ 且

$$\boldsymbol{U}_{n,\beta}(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{u}_{i,\beta}(Y_i;\boldsymbol{\theta}),$$

是 $\boldsymbol{\theta}$ 的非受限最小 *DPD* 估计量的估计函数，其中

$$\begin{aligned}
\boldsymbol{u}_{i,\beta}(y;\boldsymbol{\theta}) &= f_{i,\boldsymbol{\theta}}^{\beta}(y)\boldsymbol{s}_{i,\boldsymbol{\theta}}(y) - \boldsymbol{\xi}_{i,\beta}(\boldsymbol{\theta}),\\
\boldsymbol{s}_{i,\boldsymbol{\theta}}(y) &= \tfrac{\partial}{\partial\boldsymbol{\theta}}\log f_{i,\boldsymbol{\theta}}(y),\\
\boldsymbol{\xi}_{i,\beta}(\boldsymbol{\theta}) &= \mathrm{E}\left[\boldsymbol{u}_{i,\beta}(Y_i,\boldsymbol{\theta})\right] = \int_{-\infty}^{+\infty} f_{i,\boldsymbol{\theta}}^{\beta+1}(y)\boldsymbol{s}_{i,\boldsymbol{\theta}}(y)dy = \mathrm{E}[f_{i,\boldsymbol{\theta}}^{\beta}(Y_i)\boldsymbol{s}_{i,\boldsymbol{\theta}}(Y_i)].
\end{aligned} \tag{4}$$

所谓的离散函数 $h(\cdot)$ 从函数形式上决定了异方差的形式，但并未预先指定。假设该函数连续，至少具有一阶和二阶导数，并满足

$$h(\eta) > 0\ \forall\eta\in\mathbb{R}, \quad h(0) = 1, \quad \tfrac{d}{d\eta}h(\eta)\Big|_{\eta=0} \neq 0.$$

定义　第 2 节 条件异方差线性回归模型定义如下

$$Y_i = \boldsymbol{x}_i^T\boldsymbol{\beta} + \epsilon_i, \quad i = 1,\ldots,n, \tag{5}$$

其中 $\epsilon_i \overset{ind}{\sim} \mathcal{N}(0,\sigma_i^2)$ 且

$$\begin{aligned}
\boldsymbol{\beta} &= (\beta_0,\beta_1,\ldots,\beta_p)^T \in \mathbb{R}^{p+1},\\
\boldsymbol{x}_i^T &= (1,x_{i1},\ldots,x_{ip}) \in \mathbb{R}^{p+1}, \quad p\in\mathbb{N},\\
\boldsymbol{\alpha} &= (\alpha_1,\ldots,\alpha_r)^T \in \mathbb{R}^r,\\
\boldsymbol{z}_i^T &= (z_{i1},\ldots,z_{ir}) \in \mathbb{R}^r, \quad r\in\mathbb{N},\\
\sigma_i^2 &= \sigma_i^2(\boldsymbol{\alpha},\sigma^2) = \sigma^2 h(\boldsymbol{z}_i^T\boldsymbol{\alpha}).
\end{aligned}$$

$h(\cdot$涵盖了文献中考虑的大多数方案，例如，加性尺度模型 $h(\boldsymbol{z}_i^T\boldsymbol{\alpha}) = (1+\boldsymbol{z}_i^T\boldsymbol{\alpha})$ 或乘性尺度模型 $h(\boldsymbol{z}_i^T\boldsymbol{\alpha}) = \exp\{\boldsymbol{z}_i^T\boldsymbol{\alpha}\}$。假设解释变量 $\boldsymbol{x}_i$、 $\boldsymbol{z}_i$、 $i=1$、.....、 n均为固定的，即非随机的。用完整的矩阵表示法，式（5）可表示为

$$\boldsymbol{Y} = \mathbb{X}_n\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{6}$$

其中

$$\begin{aligned}
\boldsymbol{Y} &= (Y_1,\ldots,Y_n)^T \sim \mathcal{N}(\mathbb{X}_n\boldsymbol{\beta}, \mathrm{diag}\{\sigma_i^2(\boldsymbol{\alpha},\sigma^2)\}_{i=1}^n),\\
\mathbb{X}_n &= (\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n)^T,\\
\boldsymbol{\epsilon} &= (\epsilon_1,\ldots,\epsilon_n)^T \sim \mathcal{N}(\mathbf{0}_n, \mathrm{diag}\{\sigma_i^2(\boldsymbol{\alpha},\sigma^2)\}_{i=1}^n).
\end{aligned}$$

对于异方差线性模型（5）或（6），完整的参数向量为

$$\boldsymbol{\theta} = (\boldsymbol{\alpha}^T, \sigma^2, \boldsymbol{\beta}^T)^T, \tag{7}$$

并且，同方差性对异方差的假设检验确立了

$$H_0 : \boldsymbol{\alpha} = \mathbf{0}_r \quad vs. \quad H_1 : \boldsymbol{\alpha} \neq \mathbf{0}_r. \tag{8}$$

命题2 对于异方差线性模型（5）或（6），参数 $\boldsymbol{\theta}$ 的无约束最小 *DPD* 估计量的估计函数为 $\boldsymbol{U}_{n,\beta}(\boldsymbol{Y};\ \boldsymbol{\theta}) = \mathbf{0}_{p+r+2}$，其中

$$\boldsymbol{U}_{n,\beta}(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^n \boldsymbol{u}_{i,\beta}(Y_i;\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^n \left(\boldsymbol{u}_{i,\beta,\boldsymbol{\alpha}}^T(Y_i;\boldsymbol{\theta}), \boldsymbol{u}_{i,\beta,\sigma^2}^T(Y_i;\boldsymbol{\theta}), \boldsymbol{u}_{i,\beta,\boldsymbol{\beta}}^T(Y_i;\boldsymbol{\theta})\right)^T,$$

with

$$\boldsymbol{u}_{i,\beta,\boldsymbol{\alpha}}(Y_i;\boldsymbol{\theta}) = \frac{1}{(2\pi)^{\frac{\beta}{2}}(\sigma_i^2(\boldsymbol{\alpha},\sigma^2))^{\frac{\beta}{2}}}\frac{1}{2}\frac{h'(\boldsymbol{z}_i^T\boldsymbol{\alpha})}{h(\boldsymbol{z}_i^T\boldsymbol{\alpha})}\left[\exp\left\{-\tfrac{\beta}{2}g_i(\boldsymbol{\theta})\right\}(g_i(\boldsymbol{\theta})-1) + \frac{\beta}{(\beta+1)^{\frac{3}{2}}}\right]\boldsymbol{z}_i, \tag{9}$$

$$u_{i,\beta,\sigma^2}(Y_i;\boldsymbol{\theta}) = \frac{1}{(2\pi)^{\frac{\beta}{2}}(\sigma_i^2(\boldsymbol{\alpha},\sigma^2))^{\frac{\beta}{2}}}\frac{1}{2\sigma^2}\left[\exp\left\{-\tfrac{\beta}{2}g_i(\boldsymbol{\theta})\right\}(g_i(\boldsymbol{\theta})-1) + \frac{\beta}{(\beta+1)^{\frac{3}{2}}}\right],$$

$$\boldsymbol{u}_{i,\beta,\boldsymbol{\beta}}(Y_i;\boldsymbol{\theta}) = \frac{1}{(2\pi)^{\frac{\beta}{2}}(\sigma_i^2(\boldsymbol{\alpha},\sigma^2))^{\frac{\beta}{2}}}\exp\left\{-\tfrac{\beta}{2}g_i(\boldsymbol{\theta})\right\}\frac{1}{\sigma^2}\frac{\epsilon_i(\boldsymbol{\beta})}{h(\boldsymbol{z}_i^T\boldsymbol{\alpha})}\boldsymbol{x}_i,$$

with

$$g_i(\boldsymbol{\theta}) = \frac{\epsilon_i^2(\boldsymbol{\beta})}{\sigma_i^2(\boldsymbol{\alpha},\sigma^2)}, \tag{10}$$

$$\epsilon_i = \epsilon_i(\boldsymbol{\beta}) = Y_i - \boldsymbol{x}_i^T\boldsymbol{\beta}. \tag{11}$$

证明。根据 Martín（2021），结合 $p=1$ 和定理6，考虑参数向量 $\boldsymbol{\eta}_i=\boldsymbol{\eta}_i(\boldsymbol{\theta})=(\sigma_i^2,\ \mu_i)^T$，其中 $\mu_i=\mu_i(\boldsymbol{\beta})=\boldsymbol{x}_i^T\boldsymbol{\beta},\quad \sigma_i^2=\sigma^2h(\boldsymbol{z}_i^T\boldsymbol{\alpha})$。根据微分链式法则，可得

$$\boldsymbol{u}_i(Y_i;\boldsymbol{\theta})=\frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{\eta}_i^T(\boldsymbol{\theta})\boldsymbol{u}_i(Y_i;\boldsymbol{\eta}_{i,\beta}),$$

其中 $\boldsymbol{u}_i(Y_i;\boldsymbol{\eta}_{i,\ \beta})=(u_i(Y_i;\ \sigma_i^2)\ u_i(Y_i;\ \mu_i))^T$，

$$\frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{\eta}_i^T(\boldsymbol{\theta})=\begin{pmatrix}\sigma^2h'(\boldsymbol{z}_i^T\boldsymbol{\alpha})\boldsymbol{z}_i & 0\\ h(\boldsymbol{z}_i^T\boldsymbol{\alpha}) & 0\\ 0 & \boldsymbol{x}_i\end{pmatrix},\tag{12}$$

最后

$$\begin{aligned}\boldsymbol{u}_{i,\beta,\boldsymbol{\alpha}}(Y_i;\boldsymbol{\theta})&=\sigma^2h'(\boldsymbol{z}_i^T\boldsymbol{\alpha})u_i(Y_i;\sigma_i^2)\boldsymbol{z}_i,\\ u_{i,\beta,\sigma^2}(Y_i;\boldsymbol{\theta})&=h(\boldsymbol{z}_i^T\boldsymbol{\alpha})u_i(Y_i;\sigma_i^2),\\ \boldsymbol{u}_{i,\beta,\boldsymbol{\beta}}(Y_i;\boldsymbol{\theta})&=u_i(Y_i;\mu_i)\boldsymbol{x}_i,\end{aligned}$$

with

$$\begin{aligned}u_i(Y_i;\sigma_i^2)&=\frac{1}{(2\pi)^{\frac{\beta}{2}}(\sigma_i^2(\boldsymbol{\alpha},\sigma^2))^{\frac{\beta}{2}}}\frac{1}{2\sigma^2h(\boldsymbol{z}_i^T\boldsymbol{\alpha})}\left[\exp\left\{-\tfrac{\beta}{2}g_i(\boldsymbol{\theta})\right\}(g_i(\boldsymbol{\theta})-1)+\frac{\beta}{(\beta+1)^{\frac{3}{2}}}\right],\\ u_i(Y_i;\mu_i)&=\frac{1}{(2\pi)^{\frac{\beta}{2}}(\sigma_i^2(\boldsymbol{\alpha},\sigma^2))^{\frac{\beta}{2}}}\exp\left\{-\tfrac{\beta}{2}g_i(\boldsymbol{\theta})\right\}\frac{1}{\sigma^2}\frac{\epsilon_i(\boldsymbol{\beta})}{h(\boldsymbol{z}_i^T\boldsymbol{\alpha})}.\end{aligned}$$

■

推论3 对于异方差线性模型 *(5)* 或 *(6)*，在同方差原假设 $\boldsymbol{\theta}$ 下，受限最小*DPD*估计量 $\widetilde{\boldsymbol{\theta}}_\beta=(\mathbf{0}_r^T,\ \widetilde{\sigma}_\beta^2,\ \widetilde{\boldsymbol{\beta}}_\beta^T)^T$可通过求解以下方程得到：

$$\mathbb{X}_n^T\mathrm{diag}\left(\exp(-\tfrac{\beta}{2}\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_\beta))\right)\left(\boldsymbol{Y}-\mathbb{X}_n\widetilde{\boldsymbol{\beta}}_\beta\right)=\mathbf{0}_r,\tag{13}$$

$$\mathbf{1}_n^T\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_\beta)=0,\tag{14}$$

其中

$$\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_\beta) = (g_1(\widetilde{\boldsymbol{\theta}}_\beta), \ldots, g_n(\widetilde{\boldsymbol{\theta}}_\beta))^T,$$

$$g_i(\widetilde{\boldsymbol{\theta}}_\beta) = \frac{\epsilon_i^2(\widetilde{\boldsymbol{\beta}}_\beta)}{\widetilde{\sigma}_\beta^2}, \quad \epsilon_i(\widetilde{\boldsymbol{\beta}}_\beta) = Y_i - \boldsymbol{x}_i^T\widetilde{\boldsymbol{\beta}}_\beta, \quad i = 1, \ldots, n,$$

$$\exp(-\tfrac{\beta}{2}\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_\beta)) = (\exp\{-\tfrac{\beta}{2}g_1(\widetilde{\boldsymbol{\theta}}_\beta)\}, \ldots, \exp\{-\tfrac{\beta}{2}g_n(\widetilde{\boldsymbol{\theta}}_\beta)\})^T, \tag{15}$$

$$\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_\beta) = \operatorname{diag}\left(\exp(-\tfrac{\beta}{2}\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_\beta))\right)(\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_\beta) - \mathbf{1}_n) + \frac{\beta}{(\beta+1)^{\frac{3}{2}}}\mathbf{1}_n. \tag{16}$$

此外，*LM*向量对应的受限最小*DPD*估计量为

$$\boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta) = \frac{1}{(2\pi)^{\frac{\beta}{2}}(\widetilde{\sigma}_\beta^2)^{\frac{\beta}{2}}}\frac{h'(0)}{2}\frac{1}{n}\mathbb{Z}_n^T\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_\beta), \tag{17}$$

with

$$\mathbb{Z}_n = (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n)^T,$$

具有 r 列（n维向量），且与 $\mathbf{1}_n$线性无关。

证明。 (13)直接由命题2得出，取 $\boldsymbol{\alpha} = \mathbf{0}_r$ 和$\boldsymbol{u}_{i,\beta,\boldsymbol{\beta}}(Y_i;\boldsymbol{\theta}) = \mathbf{0}_{p+1}$，而则来自 (17)

$$\sum_{i=1}^n \boldsymbol{u}_{i,\beta,\boldsymbol{\alpha}=\mathbf{0}_p}(Y_i;\widetilde{\boldsymbol{\theta}}_\beta) = \frac{\frac{h'(0)}{2}}{(2\pi)^{\frac{\beta}{2}}(\widetilde{\sigma}_\beta^2)^{\frac{\beta}{2}}}\sum_{i=1}^n\left(\exp\left\{-\tfrac{\beta}{2}g_i(\widetilde{\boldsymbol{\theta}}_\beta)\right\}\left(g_i(\widetilde{\boldsymbol{\theta}}_\beta) - 1\right) + \frac{\beta}{(\beta+1)^{\frac{3}{2}}}\right)\boldsymbol{z}_i$$

$\sim$ n $\sim$和 $\boldsymbol{\lambda}_{n,\beta}(\boldsymbol{\theta}_\beta) = \frac{1}{n}\sum \boldsymbol{u}_{i,\beta,}$ ■

$\boldsymbol{\alpha}=\mathbf{0}_p(Y_i;\boldsymbol{\theta}_\beta)$，根据（2）。$i=1$

3 主要理论结果

本节将重点讨论第2节中引入的一般模型。对于同分布观测这一特殊情形，$\sqrt{n}(\boldsymbol{\theta}_\beta - \boldsymbol{\theta}_0)$ 的渐近分布已由 Basu 等（2017）给出。以下结果从两个方面对前述结果进行了推广：将其扩展至非同分布观测，并同时考虑拉格朗日乘子的估计量。在下文中，假设满足 Basu 等（2017）基于单个观测值给出的正则条件，且以下两个基于全体观测值的矩阵的极限存在：

$$\boldsymbol{J}_\beta(\boldsymbol{\theta}) = \lim_{n\to\infty}\bar{\boldsymbol{J}}_{n,\beta}(\boldsymbol{\theta}) = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^n \boldsymbol{J}_{i,\beta}(\boldsymbol{\theta}), \tag{18}$$

with

$$\boldsymbol{J}_{i,\beta}(\boldsymbol{\theta}) = -\mathrm{E}\left[\frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{u}_{i,\beta}^{T}\left(Y_i,\boldsymbol{\theta}\right)\right] = \int_{-\infty}^{+\infty} f_{i,\boldsymbol{\theta}}^{\beta+1}(y)\boldsymbol{s}_{i,\boldsymbol{\theta}}(y)\boldsymbol{s}_{i,\boldsymbol{\theta}}^{T}(y)dy = \mathrm{E}[f_{i,\boldsymbol{\theta}}^{\beta}(Y_i)\boldsymbol{s}_{i,\boldsymbol{\theta}}(Y_i)\boldsymbol{s}_{i,\boldsymbol{\theta}}^{T}(Y_i)],$$

and

$$\boldsymbol{K}_{\beta}\left(\boldsymbol{\theta}\right) = \lim_{n\to\infty}\bar{\boldsymbol{K}}_{n,\beta}(\boldsymbol{\theta}) = \lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{K}_{i,\beta}\left(\boldsymbol{\theta}\right), \tag{19}$$

with

$$\bar{\boldsymbol{K}}_{n,\beta}(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{K}_{i,\beta}\left(\boldsymbol{\theta}\right) = \bar{\boldsymbol{J}}_{n,2\beta}(\boldsymbol{\theta}) - \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\xi}_{i,\beta}\left(\boldsymbol{\theta}\right)\boldsymbol{\xi}_{i,\beta}^{T}\left(\boldsymbol{\theta}\right),$$
$$\boldsymbol{K}_{i,\beta}(\boldsymbol{\theta}) = \mathrm{E}\left[\boldsymbol{u}_{i,\beta}\left(Y_i,\boldsymbol{\theta}\right)\boldsymbol{u}_{i,\beta}^{T}\left(Y_i,\boldsymbol{\theta}\right)\right] = \boldsymbol{J}_{i,2\beta}(\boldsymbol{\theta}) - \boldsymbol{\xi}_{i,\beta}(\boldsymbol{\theta})\boldsymbol{\xi}_{i,\beta}^{T}(\boldsymbol{\theta}),$$

以及 $\boldsymbol{\xi}_{i,\ \beta}(\boldsymbol{\theta})$ 由（4）给出。

定理 4 在式 *(1)* 所确立的原假设下，并满足式 *(18)-(19)* 的存在性假设，有

$$\sqrt{n}\begin{bmatrix}\widetilde{\boldsymbol{\theta}}_{\beta}-\boldsymbol{\theta}_0\\ \boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_{\beta})\end{bmatrix}\xrightarrow[n\to\infty]{\mathcal{L}}\mathcal{N}\left(\mathbf{0}_p,\begin{bmatrix}\boldsymbol{P}_{\beta}(\boldsymbol{\theta}_0) & \boldsymbol{R}_{\beta}(\boldsymbol{\theta}_0)\\ \boldsymbol{R}_{\beta}^{T}(\boldsymbol{\theta}_0) & \boldsymbol{Q}_{\beta}(\boldsymbol{\theta}_0)\end{bmatrix}\right),$$

其中，正态分布的方差—协方差矩阵是奇异的，且

$$\boldsymbol{P}_{\beta}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}_{11,\beta}(\boldsymbol{\theta})\boldsymbol{K}_{\beta}(\boldsymbol{\theta})\boldsymbol{\Sigma}_{11,\beta}^{T}(\boldsymbol{\theta}), \tag{20}$$
$$\boldsymbol{Q}_{\beta}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta})\boldsymbol{K}_{\beta}(\boldsymbol{\theta})\boldsymbol{\Sigma}_{21,\beta}^{T}(\boldsymbol{\theta}), \tag{21}$$
$$\boldsymbol{R}_{\beta}(\boldsymbol{\theta}) = \boldsymbol{\Sigma}_{11,\beta}(\boldsymbol{\theta})\boldsymbol{K}_{\beta}(\boldsymbol{\theta})\boldsymbol{\Sigma}_{21,\beta}^{T}(\boldsymbol{\theta}), \tag{22}$$

and

$$\boldsymbol{\Sigma}_{11,\beta}(\boldsymbol{\theta}) = \boldsymbol{J}_{\beta}^{-1}(\boldsymbol{\theta}) - \boldsymbol{J}_{\beta}^{-1}(\boldsymbol{\theta})\boldsymbol{M}(\boldsymbol{\theta})\boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}),$$
$$\boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}) = \left[\boldsymbol{M}^{T}(\boldsymbol{\theta})\boldsymbol{J}_{\beta}^{-1}(\boldsymbol{\theta})\boldsymbol{M}(\boldsymbol{\theta})\right]^{-1}\boldsymbol{M}^{T}(\boldsymbol{\theta})\boldsymbol{J}_{\beta}^{-1}(\boldsymbol{\theta}),$$
$$.\boldsymbol{\Sigma}_{22,\beta}(\boldsymbol{\theta}) = \left[\boldsymbol{M}^{T}(\boldsymbol{\theta})\boldsymbol{J}_{\beta}^{-1}(\boldsymbol{\theta})\boldsymbol{M}(\boldsymbol{\theta})\right]^{-1}.$$

附录 A 包含定理 4的证明。Aitchison 和 Silvey（1958）以及 Silvey（1959）发表的原始经典拉格朗日乘子检验针对的是同分布¯ ¯ ¯观测，其中考虑到对于 $\beta = 0$, $\boldsymbol{K}_{n,\beta=0}(\boldsymbol{\theta}) = \boldsymbol{J}_{n,\beta=0}(\boldsymbol{\theta}) = \boldsymbol{I}_{F,n}(\boldsymbol{\theta})$ 是平均信息矩阵，这些矩阵的作用由基于单个观测的信息矩阵 ˜承担。$\sqrt{n}\boldsymbol{\lambda}_{n,\beta}(\boldsymbol{\theta}_{\beta})$ 的渐近分布是以下定义的基石，并推广了 Basu 等人（2022）给出的定义；基于该定义无法推导出本文第 β4 节介绍的 Breusch-Pagan β-分数 LM 检验。

定义3 针对非同分布个体的 β-分数拉格朗日乘子检验定义如下：

$$R_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta) = n\boldsymbol{U}_{n,\beta}^T(\widetilde{\boldsymbol{\theta}}_\beta)\bar{\boldsymbol{J}}_{n,\beta}^{-1}(\widetilde{\boldsymbol{\theta}}_\beta)\boldsymbol{M}(\widetilde{\boldsymbol{\theta}}_\beta)\left[\boldsymbol{M}^T(\widetilde{\boldsymbol{\theta}}_\beta)\bar{\boldsymbol{V}}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\boldsymbol{M}(\widetilde{\boldsymbol{\theta}}_\beta)\right]^{-1}\boldsymbol{M}^T(\widetilde{\boldsymbol{\theta}}_\beta)\bar{\boldsymbol{J}}_{n,\beta}^{-1}(\widetilde{\boldsymbol{\theta}}_\beta)\boldsymbol{U}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta), \tag{23}$$

其中

$$\bar{\boldsymbol{V}}_{n,\beta}(\boldsymbol{\theta}) = \bar{\boldsymbol{J}}_{n,\beta}^{-1}(\boldsymbol{\theta})\bar{\boldsymbol{K}}_{n,\beta}(\boldsymbol{\theta})\bar{\boldsymbol{J}}_{n,\beta}^{-1}(\boldsymbol{\theta}) \tag{24}$$

是“经验三明治矩阵”。

定理5 *(23)* 的渐近分布服从自由度为 r 的卡方分布。

检验统计量取较大值，落在 χ_r^2的右尾区域，表明有强有力的证据反对原假设。附录B 给出了定理5的证明。

定理6 考虑检验 H_0： $\boldsymbol{\theta} \in \Theta_0$ 与 H_1： $\boldsymbol{\theta} \in \Theta - \Theta_0$，其中 Θ_0 由（*9*）给出，并令局部皮特曼型备择假设序列定义为 $H_{1,\ n}$： $\boldsymbol{\theta}_n \in \Theta_{1,\ n}$，其中 $\Theta_{1,\ n}$ 由下式给出

$$\Theta_{1,n} = \left\{\boldsymbol{\theta}_n \in \Theta : \boldsymbol{m}(\boldsymbol{\theta}) = \tfrac{1}{\sqrt{n}}\boldsymbol{\delta}\right\}, \tag{25}$$

固定不变 $\boldsymbol{\delta} \in \mathbb{R}^r - \{\mathbf{0}_r\}$。在序列 $H_{1,\ n}$： $\boldsymbol{\theta} \in \Theta_{1,\ n}$下，用于非同分布个体的 β-分数 $\boldsymbol{L}$ $\boldsymbol{M}$ 检验统计量 *(23)* 的渐近分布服从自由度为 r 、非中心参数为

$$\nu_\beta(\boldsymbol{\theta}_0, \boldsymbol{\delta}) = \boldsymbol{\delta}^T\left[\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{V}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\right]^{-1}\boldsymbol{\delta}, \tag{26}$$

其中

$$\boldsymbol{V}_\beta(\boldsymbol{\theta}) = \boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta})\boldsymbol{K}_\beta(\boldsymbol{\theta})\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}) \tag{27}$$

~是“理论三明治矩阵”，即 $\boldsymbol{\theta}_\beta$ 的渐近方差—协方差矩阵。给定名义水平 α，相应的渐近功效函数为

$$\pi_\beta(\boldsymbol{y}; \boldsymbol{\theta}_0, \boldsymbol{\delta}) = \lim_{n\to\infty} P(R_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta) > \chi_{r,\alpha}^2 | \underline{F}_n(\boldsymbol{y}; \boldsymbol{\theta}_0, \boldsymbol{\delta})) = Q_{\frac{r}{2}}(\nu_\beta(\boldsymbol{\theta}_0, \boldsymbol{\delta}), \chi_{r,\alpha}^2), \tag{28}$$

其中 $Q_M(a, b)$ 是马库姆 Q 函数（详见 *Nuttal，1975* 年）。

根据前述结果，Breusch-Pagan 检验在皮特曼备择假设下是一致的，因为取 $\boldsymbol{\delta}^* = \sqrt{n}\boldsymbol{\delta}$ 后，我们得到 $\lim_{\nu\to\infty} Q_{\frac{r}{2}}(\nu, \chi_{r,\alpha}^2) = 1$。

β-分数 LM 检验相对于经典 LM 检验的皮特曼渐近相对效率（渐近相对效率）为

$$\mathrm{ARE}(R_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta), R_{n,\beta=0}(\widetilde{\boldsymbol{\theta}}_{\beta=0}); \boldsymbol{\delta}) = \frac{\nu_{\beta=0}(\boldsymbol{\theta}_0, \boldsymbol{\delta})}{\nu_\beta(\boldsymbol{\theta}_0, \boldsymbol{\delta})} > 1, \tag{29}$$

基于 Hannan（1956）以及 Koenker 和 Bassett（1982）提出的类似思路。

作为定义 3的特殊情形，我们得到以下版本。

命题 7 考虑 $\boldsymbol{\theta}=(\boldsymbol{\theta}_1^T,\boldsymbol{\theta}_2^T,\boldsymbol{\theta}_3^T)^T$和 $\boldsymbol{m}(\boldsymbol{\theta})=\boldsymbol{\theta}_1-\boldsymbol{\theta}_{1,0}$ 的特殊情形，其中 $(\widetilde{\boldsymbol{\theta}}_{1,\beta}^T,\widetilde{\boldsymbol{\theta}}_{2,\beta}^T)^T$ 与 $\widetilde{\boldsymbol{\theta}}_{3,\beta}$ 渐近独立，并且在对应于 $(\widetilde{\boldsymbol{\theta}}_{1,\beta}^T,\widetilde{\boldsymbol{\theta}}_{2,\beta}^T)^T$的区块内，相对于 $\bar{\boldsymbol{J}}_{n,\beta}(\boldsymbol{\theta})$，$\bar{\boldsymbol{K}}_{n,\beta}(\boldsymbol{\theta})$ 是一个比例矩阵。那么，式 (23) 中给出的针对非同分布个体的 β-分数拉格朗日乘子检验具有以下更简单的表达式

$$R_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)=n\boldsymbol{\lambda}_{n,\beta}^T(\widetilde{\boldsymbol{\theta}}_\beta)\left(\bar{\boldsymbol{K}}_{n,\beta}^{-1}(\widetilde{\boldsymbol{\theta}}_\beta)\right)_{11}\boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta). \tag{30}$$

其中

$$\left(\bar{\boldsymbol{K}}_{n,\beta}^{-1}(\widetilde{\boldsymbol{\theta}}_\beta)\right)_{11}=(\boldsymbol{I}_r,\mathbf{0}_{r\times(p-r)})\bar{\boldsymbol{K}}_{n,\beta}^{-1}(\boldsymbol{\theta})(\boldsymbol{I}_r,\mathbf{0}_{r\times(p-r)})^T.$$

4 Breusch-Pagan β-分数 LM 检验的推导

聚焦于定义3的矩阵形式，我们来考察异方差函数的扩展设计矩阵 $\breve{\mathbb{Z}}_n=(1_n,\mathbb{Z}_n)$ 及其对应的行 $\breve{\boldsymbol{z}}_i=(1,\boldsymbol{z}_i)^T$, $i=1,\ldots,n$。我们假设 $\Lambda_{\min}(\frac{1}{n}\mathbb{X}_n^T\mathbb{X}_n)>0$, $\Lambda_{\min}(\frac{1}{n}\breve{\mathbb{Z}}_n^T\breve{\mathbb{Z}}_n)>0$ 以及

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\frac{\max\limits_{1\le i\le n}\|\boldsymbol{x}_i\|}{\Lambda_{\min}^{\frac{1}{2}}(\frac{1}{n}\mathbb{X}_n^T\mathbb{X}_n)}=\lim_{n\to\infty}\frac{1}{\sqrt{n}}\frac{\max\limits_{1\le i\le n}\|\breve{\boldsymbol{z}}_i\|}{\Lambda_{\min}^{\frac{1}{2}}(\frac{1}{n}\breve{\mathbb{Z}}_n^T\breve{\mathbb{Z}}_n)}=0, \tag{31}$$

其中 $\Lambda_{\min}(\cdot)$ 表示矩阵的最小特征值。

使用（7）中给出的完整参数，关于同方差性与异方差性的假设（8），属于定理 7的一个特例，其中 $\boldsymbol{\alpha}$ 起到$\boldsymbol{\theta}_1$ 的作用，而 $(\sigma^2,\boldsymbol{\beta}^T)^T$起到 $\boldsymbol{\theta}_2$的作用。本节主要致力于计算 $\bar{\boldsymbol{K}}_{n,\beta,\boldsymbol{\alpha}}(\widetilde{\boldsymbol{\theta}}_\beta)$ 和 $\bar{\boldsymbol{J}}_{n,\beta,\boldsymbol{\alpha}}^{-1}(\widetilde{\boldsymbol{\theta}}_\beta)$。

定理 8 对于异方差线性模型 (5) 或 (6)，在同方差原假设下，有

$$\bar{\boldsymbol{K}}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)=\begin{pmatrix}a_{1,\beta}\boldsymbol{W}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta) & \mathbf{0}_{(r+1)\times(p+1)}\\ \mathbf{0}_{(p+1)\times(r+1)} & a_{2,\beta}\frac{1}{n}\mathbb{X}_n^T\mathbb{X}_n\end{pmatrix}, \tag{32}$$

$$\bar{\boldsymbol{J}}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)=\begin{pmatrix}b_{1,\beta}\boldsymbol{W}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta) & \mathbf{0}_{(r+1)\times(p+1)}\\ \mathbf{0}_{(p+1)\times(r+1)} & b_{2,\beta}\frac{1}{n}\mathbb{X}_n^T\mathbb{X}_n\end{pmatrix}, \tag{33}$$

$$\boldsymbol{W}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)=\begin{pmatrix}\widetilde{\sigma}_\beta^4(h'(0))^2\frac{1}{n}\mathbb{Z}_n^T\mathbb{Z}_n & \widetilde{\sigma}_\beta^2h'(0)\overline{\boldsymbol{z}}_n\\ \widetilde{\sigma}_\beta^2h'(0)\overline{\boldsymbol{z}}_n^T & 1\end{pmatrix}, \tag{34}$$

其中 $\frac{1}{n}\mathbb{X}_n^T\mathbb{X}_n=\frac{1}{n}\sum_{i=1}^n \boldsymbol{x}_i\boldsymbol{x}_i^T$, $\frac{1}{n}\mathbb{Z}_n^T\mathbb{Z}_n=\frac{1}{n}\sum_{i=1}^n \boldsymbol{z}_i\boldsymbol{z}_i^T$, $\bar{\boldsymbol{z}_n} = \frac{1}{n}\mathbb{Z}_n^T\mathbf{1}_n$,

$$a_{1,\beta} = \frac{\frac{2\beta^2+1}{2(2\beta+1)^{5/2}} - \frac{\beta^2}{4(\beta+1)^3}}{(2\pi)^\beta(\widetilde{\sigma}_\beta^2)^{\beta+2}}, \quad a_{2,\beta} = \frac{1}{(2\pi)^\beta(\widetilde{\sigma}_\beta^2)^{\beta+2}(2\beta+1)^{\frac{3}{2}}},$$

$$b_{1,\beta} = \frac{\frac{\beta^2+2}{4(\beta+1)^{5/2}}}{(2\pi)^{\frac{\beta}{2}}(\widetilde{\sigma}_\beta^2)^{\frac{\beta}{2}+2}}, \quad b_{2,\beta} = \frac{1}{(2\pi)^{\frac{\beta}{2}}(\widetilde{\sigma}_\beta^2)^{\frac{\beta}{2}+1}(\beta+1)^{\frac{3}{2}}}.$$

证明。 由（12），取 $h(0) = 1$，可得

$$\begin{aligned}\boldsymbol{K}_{i,\beta}(\widetilde{\boldsymbol{\theta}}_\beta) &= \mathrm{E}[\boldsymbol{u}_{i,\boldsymbol{\theta}}(Y_i;\widetilde{\boldsymbol{\theta}}_\beta)\boldsymbol{u}_{i,\boldsymbol{\theta}}^T(Y_i;\widetilde{\boldsymbol{\theta}}_\beta)] \\ &= \frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{\eta}_i^T(\boldsymbol{\theta})\Big]_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}_\beta} \boldsymbol{K}_\beta(\boldsymbol{\eta}_i(\widetilde{\boldsymbol{\theta}}_\beta)) \frac{\partial}{\partial\boldsymbol{\theta}^T}\boldsymbol{\eta}_i(\boldsymbol{\theta})\Big]_{\boldsymbol{\theta}=\widetilde{\boldsymbol{\theta}}_\beta}, \end{aligned} \tag{35}$$

其中

$$\boldsymbol{K}_\beta(\boldsymbol{\eta}_i(\widetilde{\boldsymbol{\theta}}_\beta)) = \frac{1}{(2\pi)^\beta(\widetilde{\sigma}_\beta^2)^{\beta+2}} \begin{pmatrix} \frac{2\beta^2+1}{2(2\beta+1)^{5/2}} - \frac{\beta^2}{4(\beta+1)^3} & 0 \\ 0 & \frac{\widetilde{\sigma}_\beta^2}{(2\beta+1)^{\frac{3}{2}}} \end{pmatrix}, \tag{36}$$

$\sim$ n $\sim$可由 Martín 2020，推论4得到。最后，由 $\bar{\boldsymbol{K}}_\beta(\boldsymbol{\theta}_\beta) = \frac{1}{n}\sum \boldsymbol{K}_{i,\beta}$ $(\boldsymbol{\theta}_\beta)$，得到所需的（32）表达式。（33）的表达式可用类似方法得到。 ■

备注1 （31）中的假设源于应用多元 *Lindeberg* 中心极限定理（*CLT*），该定理适用于命题2中通过求解估计方程所得的估计量；对误差项的假设更弱，仅要求均值为零且具有有限方差。

定理9 异方差线性模型（5）或（6）的 *Breusch-Pagan* β-分数 *LM* 检验统计量（相对于经典检验统计量 $\beta = 0$）的 *Pitman* 渐近相对效率为

$$\mathrm{ARE}(R_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta), R_{n,\beta=0}(\widetilde{\boldsymbol{\theta}}_{\beta=0})) = \frac{8(\beta+1)^5}{(\beta^2+2)^2}\left(\frac{2\beta^2+1}{2(2\beta+1)^{5/2}} - \frac{\beta^2}{4(\beta+1)^3}\right), \quad \beta > 0. \tag{37}$$

Proof. 由定理8和公式（29），可得渐近相对效率为$2b_{1,\beta}^2/a_{1,\beta}$。 ■

对于纯正态分布数据，式（37）仅取决于 β（与 α、 $\boldsymbol{\theta}_0$ 或 $\boldsymbol{\delta}$无关），且实际上随 β 增加而增大（见图1）。这意味着，与最小 DPD 估计量通常的情况一样，检验功效最高的检验统计量是 β 的取值最接近经典检验统计量 $\beta = 0$的那些。可以观察到，在 $\beta = 0.2$处，渐近相对效率（ARE）为$\widetilde{}$1.09；因此，对于 $\beta = 0.2$， $R_{n,\beta=0.2}(\boldsymbol{\theta}_{\beta=0.2})$需要增加约9%的观测值，才能获得与$R_{n,\beta=0}(\boldsymbol{\theta}_{\beta=0})$相同的检验功效。

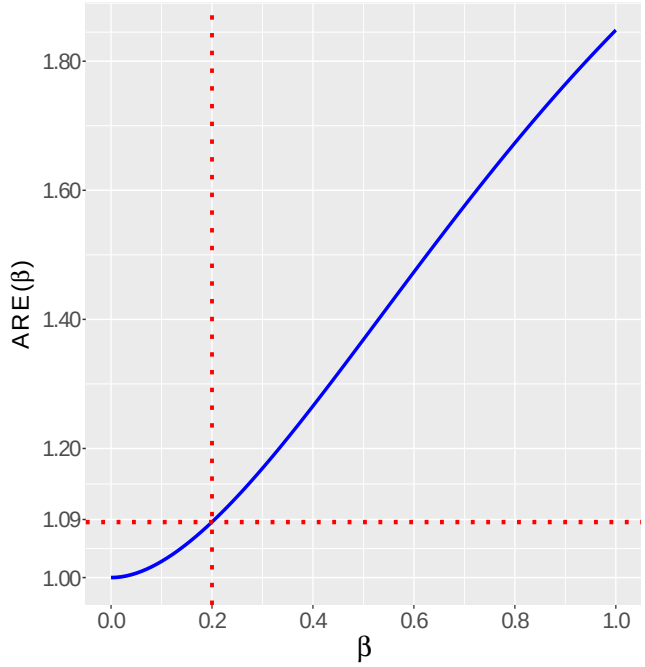


图1：渐近相对 效率：Breusch-Pagan β-分数 LM 检验统计量 .

推论10 异方差线性模型的 *Breusch-Pagan* β-分数 *LM* 检验统计量，见式（*5*）或（*6*），其表达式为

$$R_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta) = \frac{1}{\frac{2(2\beta^2+1)}{(2\beta+1)^{5/2}} - \frac{\beta^2}{(\beta+1)^3}} \boldsymbol{v}^T(\widetilde{\boldsymbol{\theta}}_\beta)\breve{\mathbb{Z}}_n(\breve{\mathbb{Z}}_n^T\breve{\mathbb{Z}}_n)^{-1}\breve{\mathbb{Z}}_n^T\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_\beta), \tag{38}$$

其中 $\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_\beta)$ 由式（*16*）给出。一旦验证了式（*31*）所确立的两个条件，检验统计量（*38*）在渐近意义下服从一个具有 r 个自由度的卡方随机变量（同方差原假设下）。

证明。 根据定理 7，我们可以使用（30）而不是（23）。待求逆的二次型矩阵需要先进行分块矩阵求逆，并选取其 $r\times r$左上分块，即

$$\left(\bar{\boldsymbol{K}}^{-1}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\right)_{11} = \frac{1}{\frac{2\beta^2+1}{2(2\beta+1)^{5/2}} - \frac{\beta^2}{4(\beta+1)^3}} \boldsymbol{W}^{-1}_{n,\beta,11}(\widetilde{\boldsymbol{\theta}}_\beta),$$

with

$$\boldsymbol{W}^{-1}_{n,\beta,11}(\widetilde{\boldsymbol{\theta}}_\beta) = \frac{n}{\widetilde{\sigma}^4(h'(0))^2}\left(\mathbb{Z}_n^T\boldsymbol{H}_n\mathbb{Z}_n\right)^{-1},$$
$$\boldsymbol{H}_n = \boldsymbol{I}_n - \tfrac{1}{n}\mathbf{1}_n\mathbf{1}_n^T,$$

根据 Lu 和 Shiou（2002，定理 2.1.-(ii)）。由（16）和（17），我们有

$$
\begin{aligned}
R_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta) &= \frac{1}{4\left(\frac{2\beta^2+1}{2(2\beta+1)^{5/2}}-\frac{\beta^2}{4(\beta+1)^3}\right)}\boldsymbol{v}^T(\widetilde{\boldsymbol{\theta}}_\beta)\mathbb{Z}_n\left(\mathbb{Z}_n^T\boldsymbol{H}_n\mathbb{Z}_n\right)^{-1}\mathbb{Z}_n^T\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_\beta) \\
&= \frac{1}{4\left(\frac{2\beta^2+1}{2(2\beta+1)^{5/2}}-\frac{\beta^2}{4(\beta+1)^3}\right)}\boldsymbol{v}^T(\widetilde{\boldsymbol{\theta}}_\beta)\left(\boldsymbol{H}_n\mathbb{Z}_n\left(\mathbb{Z}_n^T\boldsymbol{H}_n\mathbb{Z}_n\right)^{-1}\mathbb{Z}_n^T\boldsymbol{H}_n+\tfrac{1}{n}\mathbf{1}_n\mathbf{1}_n^T\right)\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_\beta).
\end{aligned}
$$

而最终表达式（38）可由 Rao 等人（2008，第 503 页）的定理 A.45 推得。 ■

注 2 经典 *Breusch-Pagan LM* 检验统计量（$\beta = 0$），

$$
R_n(\widetilde{\boldsymbol{\theta}}) = \frac{1}{2}(\boldsymbol{g}(\widetilde{\boldsymbol{\theta}})-\mathbf{1}_r)^T\breve{\mathbb{Z}}_n(\breve{\mathbb{Z}}_n^T\breve{\mathbb{Z}}_n)^{-1}\breve{\mathbb{Z}}_n^T(\boldsymbol{g}(\widetilde{\boldsymbol{\theta}})-\mathbf{1}_r), \tag{39}
$$

与 *Godfrey*（*1989*，第 *128* 页）的式（*5.17*）一致。该表达式说明了计算经典 *Breusch-Pagan LM* 检验统计量所采用的两步普通最小二乘法（*OLS*）回归程序是合理的。对这一新方法的最一般解释也很容易得出。*Breusch-Pagan* β-分数 *LM* 检验统计量的˜值基于对 $\boldsymbol{v}(\boldsymbol{\theta}_\beta)$ 的投影； $\boldsymbol{v}(\boldsymbol{\theta}_\beta)$ 是一个向量˜，由标准化 β 残差的平方向量构成，即 $\boldsymbol{g}(\boldsymbol{\theta}_\beta)$。在该线性回归模型中，以 $\boldsymbol{Y}$ 为响应变量进行调整，并以 $\mathbb{X}_n$ 为设计矩阵。当 $(r+1)+(p+1)<n$ 时，该投影位于由 $\mathbb{Z}_n$ 的列生成的、维数为 $r+1$ 的˘向量空间上。实际操作步骤如下：“（*38*）的值可通过以下方式计算：对响应变量 $\boldsymbol{v}(\boldsymbol{\theta}_\beta)$ 关于解释变量矩阵 $\breve{\mathbb{Z}}_n$ 进行 *OLS* 回归，取其解释˜平方和，再乘以 $(^{2}(2\beta^2+1)(2\beta+1)^{5/2}-\beta^2(\beta+1)^3)^{-1}$。”

5 影响函数分析

影响函数方法由 Hampel（1968、1974）针对估计量提出，其本质上基于泛函的 Gateaux 导数。Hampel 等人（1986）对其进行了完整阐述。由于该方法既可用于分析估计量的稳健性，也可用于分析检验统计量的稳健性，因此它是分析统计程序稳健性最重要的工具。

本节将介绍适用于非同分布观测的 β-分数 LM 检验统计量的二阶影响函数。我们将按以下步骤进行：首先，在原假设下计算 LM 的最小 DPD 估计量的影响函数，然后基于此得到 β-分数 LM 检验统计量的影响函数。要充分证明该检验统计量的稳健性，原则上仅分析原始影响函数还不够，还需要进一步证明数据污染下 β-分数 LM 检验统计量的显著性水平和检验功效的稳定性。

~令 $\boldsymbol{T}_{n,\ \underline{\beta}}(G)$ 表示与 $\boldsymbol{\theta}_2$ 相关的泛函，其中 $G=(G_1,\ \underline{.....},\ G_n)^T$ 是样本中逐个观测对应的真实分布向量，$Y_i,\ \ i=1,\ .....,\ n$。在真实分布属于与同方差线性回归相关的参数模型这一假设下，同一向量记为 $F_{\boldsymbol{\theta}_0}=(F_{1,\ \boldsymbol{\theta}_0},\ .....,\ F_{n,\ \boldsymbol{\theta}_0})^T$。

定理 11 针对非同分布个体的 β-分数 *LM* 检验，其通常的一阶影响函数等于零；其二阶影响函数，或 $\boldsymbol{\lambda}_{n,\beta}(\boldsymbol{\theta}_\beta)$ 的自标准化影响函数，定义为

$$\mathcal{IF}^{(2)}\left(\boldsymbol{y}, R_{n,\beta}\left(\boldsymbol{\theta}_0\right), \underline{F_{\boldsymbol{\theta}_0}}\right)=\frac{1}{n}\sum_{i=1}^{n}\mathcal{IF}_i^{(2)}\left(y_i, R_{n,\beta}\left(\boldsymbol{\theta}_0\right), \underline{F_{\boldsymbol{\theta}_0}}\right), \tag{40}$$

其中

$$\begin{aligned}&\mathcal{IF}_i^{(2)}\left(y_i, R_{n,\beta}\left(\boldsymbol{\theta}_0\right), \underline{F_{\boldsymbol{\theta}_0}}\right) \\ &=\boldsymbol{u}_{i,\beta}^T\left(y_i;\boldsymbol{\theta}\right)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta})\boldsymbol{M}(\boldsymbol{\theta}_0)\left[\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{V}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\right]^{-1}\boldsymbol{M}^T(\boldsymbol{\theta})\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta})\boldsymbol{u}_{i,\beta}\left(y_i;\boldsymbol{\theta}\right).\end{aligned} \tag{41}$$

考虑到 $\boldsymbol{u}_{i,\beta}(y_i;\ \boldsymbol{\theta}_0)$ 有界，对任意 $\beta>0$，则对于任意 $\beta>0$, $\mathcal{IF}^{(2)}\left(\boldsymbol{y}, R_{n,\beta}(\boldsymbol{\theta}_0),\right.$ $\left.F_{\boldsymbol{\theta}_0}\right)$ 关于第 i 个观测值有界，其中$i=1,\ldots,n$；而且 $\beta>0$ 的值越大，离群值受到的降权效应就越强。对于经典得分检验（$\beta=0$），$\mathcal{IF}^{(2)}\left(\boldsymbol{y}, R_{n,\beta=0}(\boldsymbol{\theta}_0)\ \ F_{\boldsymbol{\theta}_0}\right)$ 无界，因为 $\boldsymbol{u}_{i,\beta=0}\left(y_i;\boldsymbol{\theta}_0\right)=\boldsymbol{s}_{i,\boldsymbol{\theta}_0}(y_i)$ 无界。所有观测值的影响函数有界，当且仅当与个体对相关联的影响函数全部有界；也就是说，所有观测值的影响函数有界，当且仅当 $\beta>0$。非同分布观测样本中 $R_{n,\beta}(\boldsymbol{\theta}_0)$ 的粗差敏感度（GES）定义为

$$\mathcal{GES}(R_{n,\beta}\left(\boldsymbol{\theta}_0\right))=\sup_{i\in\{1,\ldots,n\}}\mathcal{GES}_i(R_{n,\beta}\left(\boldsymbol{\theta}_0\right))=\sup_{i\in\{1,\ldots,n\}}\sup_{y_i\in\mathbb{R}}\mathcal{IF}_i^{(2)}\left(y_i, R_{n,\beta}\left(\boldsymbol{\theta}_0\right), \underline{F_{\boldsymbol{\theta}_0}}\right),$$

其中，第 i 个个体的 $R_{n,\beta}\left(\boldsymbol{\theta}_0\right)$ 的粗差敏感度（GES）$\mathcal{GES}_i(R_{n,\beta}\left(\boldsymbol{\theta}_0\right))$，等于 $\boldsymbol{\lambda}_{n,\beta}(\boldsymbol{\theta}_\beta)$ 的所谓自标准化粗差敏感度的平方 $\boldsymbol{m}(\boldsymbol{\theta}_\beta)$，或 $\boldsymbol{\theta}_\beta$。有限（无限）的粗差敏感度意味着 $R_{n,\beta}(\boldsymbol{\theta}_0)$ 的二阶影响函数有界（无界）。不同的参数化会产生 $R_{n,\beta=0}(\boldsymbol{\theta}_0)$、自标准化影响函数和自标准化粗差敏感度的唯一值（后两者首次由 Krasker 和 Welsch，1982 引入）。这就是 Rao 型检验统计量所谓的不变性，而其他常用检验统计量（如似然比和 Wald 型检验）则不具备这一性质。

推论 12 在 $\boldsymbol{\theta}=(\boldsymbol{\theta}_1^T,\boldsymbol{\theta}_2^T)^T$和 $\Theta_0=\{\boldsymbol{\theta}\in\Theta:\boldsymbol{\theta}_1=\boldsymbol{\theta}_{1,0}\}$ 的特殊情形下，β-分数非同分布个体 *LM* 检验（见（*23*））具有二阶影响函数（*40*），其中

$$\mathcal{IF}_i^{(2)}\left(y_i, R_{n,\beta}\left(\boldsymbol{\theta}_0\right), \underline{F_{\boldsymbol{\theta}_0}}\right)=\boldsymbol{u}_{i,\beta,1}^T(\boldsymbol{\theta}_0)\left(\boldsymbol{K}_{n,\beta}^{-1}(\boldsymbol{\theta}_0)\right)_{11}\boldsymbol{u}_{i,\beta,1}(\boldsymbol{\theta}_0), \tag{42}$$

and

$$\begin{aligned}\boldsymbol{u}_{i,\beta,1}\left(y_i;\boldsymbol{\theta}\right)&=\left(\boldsymbol{I}_r,\mathbf{0}_{r\times(p-r)}\right)\boldsymbol{u}_{i,\beta}\left(y_i;\boldsymbol{\theta}\right),\\ \left(\boldsymbol{K}_{n,\beta}^{-1}(\boldsymbol{\theta})\right)_{11}&=\left(\boldsymbol{I}_r,\mathbf{0}_{r\times(p-r)}\right)\boldsymbol{K}_{n,\beta}^{-1}(\boldsymbol{\theta})\left(\boldsymbol{I}_r,\mathbf{0}_{r\times(p-r)}\right)^T.\end{aligned}$$

基于 $F_{\boldsymbol{\theta}_{0,n}(\boldsymbol{\delta})}(\boldsymbol{y})$，在 $\boldsymbol{\theta}_{0,n}(\boldsymbol{\delta})\in\Theta_{1,n}$ 条件下，我们将考虑一列局部皮特曼型备择假设下的 ϵ-污染分布函数，相对于退化分布函数 $\Lambda_{\boldsymbol{y}}$，

$$\underline{F}_{\boldsymbol{n},\boldsymbol{\epsilon}}(\boldsymbol{y};\boldsymbol{\theta}_0,\boldsymbol{\delta})=\left(1-\frac{\epsilon}{\sqrt{n}}\right)\underline{F}_{\boldsymbol{\theta}_{0,n}(\boldsymbol{\delta})}(\boldsymbol{y})+\frac{\epsilon}{\sqrt{n}}\Lambda_{\boldsymbol{y}}.$$

定理 13 对于 ϵ-污染分布函数，在一列局部 *Pitman* 型备择假设 $H_{1,\ n}:\ \boldsymbol{\theta}\in\Theta_{1,\ n}$，其中 $\Theta_{1,\ n}$ 由式（25）给出，针对非同分布个体的 β-分数 *LM* 检验统计量 $R_{n,\ \beta}(\boldsymbol{\theta}_{\beta,\ \epsilon})$渐近服从自由度为 r 、非中心参数为

$$\nu_\beta(\boldsymbol{\theta}_0,\boldsymbol{\delta}_{\beta,\epsilon})=\nu_\beta(\boldsymbol{\theta}_0,\boldsymbol{\delta})+2\epsilon\boldsymbol{\delta}^T\mathcal{IF}\left(\boldsymbol{y},\boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{\theta}_0\right),\underline{F}_{\boldsymbol{\theta}_0}\right)+\epsilon^2\mathcal{IF}^{(2)}\left(\boldsymbol{y},R_{n,\beta}\left(\boldsymbol{\theta}_0\right),\underline{F}_{\boldsymbol{\theta}_0}\right),\tag{43}$$

其中 $\nu_\beta(\boldsymbol{\theta}_0,\ \boldsymbol{\delta})$ 由式（26）给出，

$$\begin{aligned}\mathcal{IF}\left(\boldsymbol{y},\boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{\theta}_0\right),\underline{F}_{\boldsymbol{\theta}_0}\right)&=\sum_{i=1}^{n}\mathcal{IF}\left(y_i,\boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{\theta}_0\right),\underline{F}_{\boldsymbol{\theta}_0}\right),\\ \mathcal{IF}\left(y_i,\boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{\theta}_0\right),\underline{F}_{\boldsymbol{\theta}_0}\right)&=\left[\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{V}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\right]^{-1}\mathcal{IF}\left(y_i,\boldsymbol{m}\left(\boldsymbol{\theta}_0\right),\underline{F}_{\boldsymbol{\theta}_0}\right),\\ \mathcal{IF}\left(y_i,\boldsymbol{m}\left(\boldsymbol{\theta}_0\right),\underline{F}_{\boldsymbol{\theta}_0}\right)&=\boldsymbol{M}^T(\boldsymbol{\theta}_0)\mathcal{IF}\left(y_i,\boldsymbol{\theta}_0,\underline{F}_{\boldsymbol{\theta}_0}\right),\\ \mathcal{IF}\left(y_i,\boldsymbol{\theta}_0,\underline{F}_{\boldsymbol{\theta}_0}\right)&=\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\boldsymbol{u}_{i,\beta}\left(y_i;\boldsymbol{\theta}_0\right),\end{aligned}\tag{44}$$

$\mathcal{IF}^{(2)}\left(\boldsymbol{y},\ R_{n,\ \beta}(\boldsymbol{\theta}_0)\ \underline{F}_{\boldsymbol{\theta}_0}\right)$ 由式（40）。此外，对应的渐近功效函数在名义水平 α下为

$$\pi_{\boldsymbol{\epsilon},\beta}(\boldsymbol{y};\boldsymbol{\theta}_0,\boldsymbol{\delta})=\lim_{n\to\infty}P(R_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)>\chi^2_{r,\alpha}|\underline{F}_{\boldsymbol{n},\boldsymbol{\epsilon}}(\boldsymbol{y};\boldsymbol{\theta}_0,\boldsymbol{\delta}))=Q_{\frac{r}{2}}(\nu_\beta(\boldsymbol{\theta}_0,\boldsymbol{\delta}_{\beta,\epsilon}),\chi^2_{r,\alpha}),\tag{45}$$

其中 $Q_M(a,\ b)$ 是 *Marcum* Q-函数。

（渐近）检验功效影响函数定义为

$$\mathcal{PIF}_\beta(\boldsymbol{y};\boldsymbol{\theta}_0,\boldsymbol{\delta})=\frac{d}{d\epsilon}\pi_{\boldsymbol{\epsilon},\beta}(\boldsymbol{y};\boldsymbol{\theta}_0,\boldsymbol{\delta})\Big|_{\epsilon=0}.$$

定理 14 对于 ϵ-污染分布函数，在一系列局部 *Pitman* 型备择假设 $H_{1,n}$: $\boldsymbol{\theta} \in \Theta_{1,n}$，其中 $\Theta_{1,n}$ 由（25）给出，β-分数 *LM* 检验统计量针对非同分布个体的检验功效影响函数，$R_{n,\beta}(\boldsymbol{\theta}_{\beta,\epsilon})$，在名义水平 α 下，定义为

$$\begin{aligned}\mathcal{PIF}_{\beta}(\boldsymbol{y};\boldsymbol{\theta}_0,\boldsymbol{\delta}) = 2\nu_{\beta}(\boldsymbol{\theta}_0,\boldsymbol{\delta})\boldsymbol{\delta}^T\mathcal{IF}\left(\boldsymbol{y},\boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{\theta}_0\right),\underline{F}_{\boldsymbol{\theta}_0}\right)\\ \times\left[Q_{\frac{r}{2}+1}(\nu_{\beta}(\boldsymbol{\theta}_0,\boldsymbol{\delta}),\chi^2_{r,\alpha}) - Q_{\frac{r}{2}}(\nu_{\beta}(\boldsymbol{\theta}_0,\boldsymbol{\delta}),\chi^2_{r,\alpha})\right],\end{aligned}$$

其中 $\mathcal{IF}\left(\boldsymbol{y},\ \boldsymbol{\lambda}_{n,\beta}(\boldsymbol{\theta}_0)\ \underline{F}_{\boldsymbol{\theta}_0}\right)$ 由（44）给出。

我们可以观察到，功效影响函数有界等价于参数 $\boldsymbol{\theta}$ 的最小密度幂散度的影响函数有界。事实上，众所周知，对于$\beta > 0$，影响函数有界；而对于 $\beta = 0$，影响函数无界。

取 $\boldsymbol{\delta} = \mathbf{0}_r$ 和 $\boldsymbol{\theta} = \boldsymbol{\theta}_0$，由于 $\boldsymbol{\theta}_n(\mathbf{0}_r) = \boldsymbol{\theta}_0 \in \Theta_0$，基于 $F_{\boldsymbol{\theta}_0}(\boldsymbol{y})$，我们将考虑$F_{\boldsymbol{n},\boldsymbol{\epsilon}}(\boldsymbol{y};\mathbf{0}_r)$ 为原假设下受 ϵ 污染的分布函数，并且

$$\alpha_{\boldsymbol{\epsilon},\beta}(\boldsymbol{y}) = \pi_{\boldsymbol{\epsilon},\beta}(\boldsymbol{y};\boldsymbol{\theta}_0,\mathbf{0}_r) = \lim_{n\to\infty} P(R_{n,\beta}(\widetilde{\boldsymbol{\theta}}_{\beta}) > \chi^2_{r,\alpha}|\underline{F}_{\boldsymbol{n},\boldsymbol{\epsilon}}(\boldsymbol{y};\boldsymbol{\theta}_0)).$$

在这种情况下，$\mathcal{LIF}_{\beta}(\boldsymbol{y}) = \frac{d}{d\epsilon}\alpha_{\boldsymbol{\epsilon},\beta}(\boldsymbol{y})\big|_{\epsilon=0} = 0$，这表明显著性水平的影响函数为零。

例 15 对于 *Breusch-Pagan* β-分数 *LM* 检验统计量，二阶影响函数为

$$\mathcal{IF}^{(2)}\left(\boldsymbol{y},R_{n,\beta}\left(\mathbf{0}_r,\sigma_0^2,\boldsymbol{\beta}_0\right),\underline{F}_{\mathbf{0}_r,\sigma_0^2,\boldsymbol{\beta}_0}\right) = \frac{1}{n}\sum_{i=1}^{n}\mathcal{IF}_i^{(2)}\left(y_i,R_{n,\beta}\left(\mathbf{0}_r,\sigma_0^2,\boldsymbol{\beta}_0\right),\underline{F}_{\mathbf{0}_r,\sigma_0^2,\boldsymbol{\beta}_0}\right),$$

其中

$$\mathcal{IF}_i^{(2)}\left(y_i,R_{n,\beta}\left(\mathbf{0}_r,\sigma_0^2,\boldsymbol{\beta}_0\right),\underline{F}_{\mathbf{0}_r,\sigma_0^2,\boldsymbol{\beta}_0}\right) = \frac{\exp\left\{-\beta\frac{(y_i-\boldsymbol{x}_i^T\boldsymbol{\beta}_0)^2}{\sigma_0^2}\right\}\left(\frac{(y_i-\boldsymbol{x}_i^T\boldsymbol{\beta}_0)^2}{\sigma_0^2}-1\right)^2}{4\left(\frac{2\beta^2+1}{2(2\beta+1)^{5/2}}-\frac{\beta^2}{4(\beta+1)^3}\right)}\boldsymbol{z}_i^T\left(\mathbb{Z}_n^T\mathbb{Z}_n\right)^{-1}\boldsymbol{z}_i.$$

对于任意 $\beta > 0$, $\mathcal{IF}_i^{(2)}(y, R_{n,\beta}(\sigma_0^2,\boldsymbol{\beta}_0))$，相对于 y有界；但对于经典 *Breusch-Pagan* 得分 *LM* 检验（$\beta = 0$），则无界。此外，杠杆值不会对 *Breusch-Pagan* β-分数 *LM* 检验的二阶影响函数产生负面影响，因为第一个因子相对于 $\boldsymbol{x}_i$ 有界，第二个因子相对于 $\boldsymbol{z}_i$也有界。特别是，对于简单线性回归，在 $p = 1$, $x_{i1} = z_{i1} = i$ *((31)* 得到验证*)*的情况下，$\beta_0 = 0$, $\beta_1 = 1$, $\alpha_1 = 0$，其中 $G = F_{(0,\sigma^2,0,1)}$ 是真实模型，我们得到

$$\mathcal{IF}_i^{(2)}(y_i,R_{n,\beta}\left(0,\sigma^2,0,1\right),\underline{F}_{(0,\sigma^2,0,1)}) = \frac{\exp\left\{-\beta\frac{(y_i-i)^2}{\sigma^2}\right\}\left(\frac{(y_i-i)^2}{\sigma^2}-1\right)^2}{2\left(\frac{2\beta^2+1}{2(2\beta+1)^{5/2}}-\frac{\beta^2}{4(\beta+1)^3}\right)}\frac{3i^2}{(2n+1)(n^2-1)},$$

图 2绘出了在 $i=1$, $\sigma^2 \in \{2,6\}$和 $\beta \in \{0, 0.15, 0.2, 0.3\}$时的二阶影响函数。对于与 $\beta > 0$相关的曲线，其左右尾部清晰地呈现出降权效应；而对于 $\beta = 0$ ，曲线则无限增大。曲线的最大值出现在 y_i 处的 $(y_i - i)^2 = {}^{\beta+2}_{\beta}\sigma^2$的一个二重解处，据此得出

$$\mathcal{GES}(R_{n,\beta}\left(0,\sigma^2,0,1\right)) = 6\left[\beta^2 \exp\left\{\beta+2\right\}\left(\frac{2\beta^2+1}{2(2\beta+1)^{5/2}} - \frac{\beta^2}{4(\beta+1)^3}\right)\right]^{-1}\frac{n^2}{(2n+1)(n^2-1)}, \quad (46)$$

在 β 上递减，且仅在 $\beta > 0$ 时有限。由于 $\mathcal{GES}(R_{n,\beta}(0, \sigma^2, 0, 1)) = \infty$ 对于$\beta = 0$，经典 *Breusch-Pagan LM* 检验统计量的稳定性在存在离群值时完全崩溃，而 *Breusch-Pagan* β-分数 *LM* 检验统计量的稳健性会随着 $\beta > 0$ 的增加而增强（粗差敏感度降低）。

6 Koenker 的 β-分数检验统计量

Koenker (1981) 扩展了 Breusch-Pagan 检验统计量，使其也适用于误差项不服从正态分布的情况。通过将原始 Breusch-Pagan 检验统计量乘以一个常数，研究发现新方案具有相同的渐近分布，同时赋予其新的普通最小二乘法解释，并揭示了其与样本峰度之间的关系。同样，我们将扩展 Breusch-Paganβ-分数 LM 检验，并将其称为 Koenker 的 β-分数检验。对于经典 Breusch-Pagan 检验（$\beta = 0$），估计函数交叉乘积的平均值为

$$\tfrac{1}{n}\boldsymbol{v}^T(\widetilde{\boldsymbol{\theta}}_{\beta=0})\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_{\beta=0}) = \tfrac{1}{n}(\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_{\beta=0}) - \mathbf{1}_r)^T(\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_{\beta=0}) - \mathbf{1}_r) = \tfrac{1}{n}\boldsymbol{g}^T(\widetilde{\boldsymbol{\theta}}_{\beta=0})\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_{\beta=0}) - 1$$

其中 $\frac{1}{n}\boldsymbol{g}^T(\boldsymbol{\theta}_{\beta=0})\boldsymbol{g}(\boldsymbol{\theta}_{\beta=0}$ 是i第 - 个误差项（在同方差性原假设下）理论峰度系数的一致估计量，即 $E\left[g_i^2(\boldsymbol{\theta})\right] = {}^{E[\epsilon_i^4(\ \boldsymbol{\beta}\ ])}_{\sigma^4}$。在真正的线性同方差回归中，即当 $\epsilon_i(\boldsymbol{\beta}) \sim \mathcal{N}(0,\ \sigma^2)$ 时，理论峰度系数为 3。因此，很容易看出，若在（39）的分母中用 $\frac{1}{n}\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_{\beta=0})\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_{\beta=0})$ 替换 2，就得到 Koenker 检验统计量

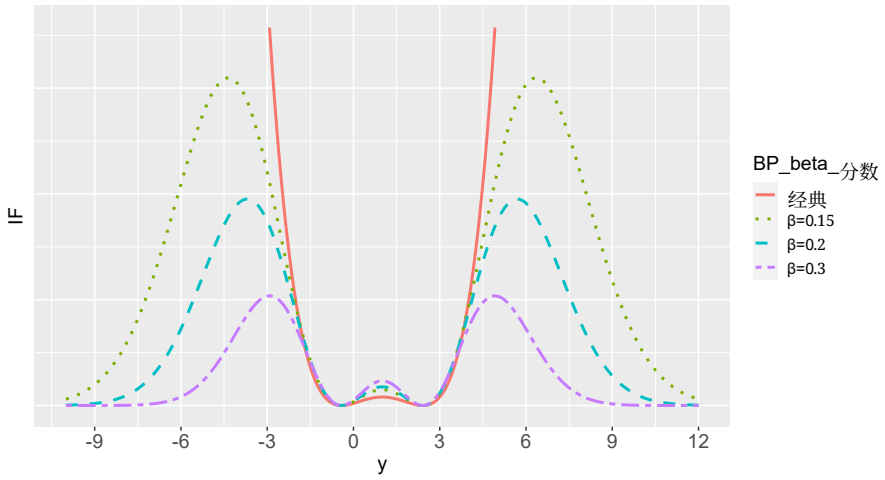
$$Q_n(\widetilde{\boldsymbol{\theta}}_{\beta=0}) = \frac{1}{\frac{1}{n}\boldsymbol{g}^T(\widetilde{\boldsymbol{\theta}}_{\beta=0})\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_{\beta=0}) - 1}(\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_{\beta=0}) - \mathbf{1}_r)^T\breve{\mathbb{Z}}_n(\breve{\mathbb{Z}}_n^T\breve{\mathbb{Z}}_n)^{-1}\breve{\mathbb{Z}}_n^T(\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_{\beta=0}) - \mathbf{1}_r). \quad (47)$$

同理，Koenker 的 β-分数检验统计量定义为

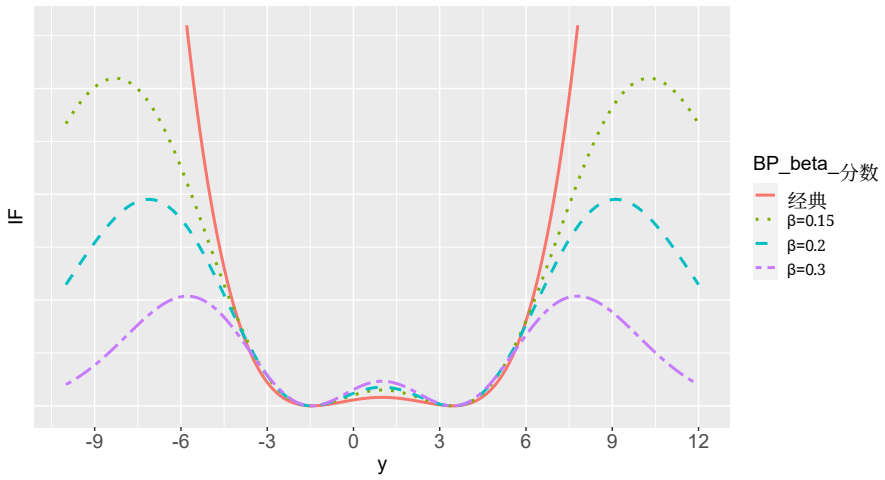
$$Q_n(\widetilde{\boldsymbol{\theta}}_{\beta}) = \frac{1}{\frac{1}{n}\boldsymbol{v}^T(\widetilde{\boldsymbol{\theta}}_{\beta})\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_{\beta})}\boldsymbol{v}^T(\widetilde{\boldsymbol{\theta}}_{\beta})\breve{\mathbb{Z}}_n(\breve{\mathbb{Z}}_n^T\breve{\mathbb{Z}}_n)^{-1}\breve{\mathbb{Z}}_n^T\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_{\beta}), \quad (48)$$

其中

$$\tfrac{1}{n}\boldsymbol{v}^T(\widetilde{\boldsymbol{\theta}}_{\beta})\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_{\beta}) = \tfrac{1}{n}\left(\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_{\beta}) - \mathbf{1}_r\right)^T \operatorname{diag}\left(\exp(-\beta\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_{\beta}))\right)\left(\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_{\beta}) - \mathbf{1}_r\right) - \frac{\beta^2}{(\beta+1)^3}.$$



(a) 情形： $\sigma^2 = 2$



(b) 情形： $\sigma^2 = 6$

图2： $\mathcal{IF}^{(2)}_{i=1}(y_{i=1},\ R_{n,\ \beta}(0,\ \sigma^2,\ 0,\ \underline{1})\ F_{(0,\ \sigma^2,\ 0,\ 1)})$ 简单线性回归的 Breusch-Pagan β-分数 LM 检验。

推论16 *Koenker* 的 β-分数检验统计量适用于异方差线性模型（*5*）或（*6*），由（*48*）给出，其渐近服从自由度为 r 的卡方分布。

证明。根据大数弱定律（WLLN），并考虑到$\widetilde{\boldsymbol{\theta}}_\beta \xrightarrow[n\to\infty]{p} \boldsymbol{\theta}$，可得

$$\tfrac{1}{n}\left(\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_\beta)-\mathbf{1}_r\right)^T \operatorname{diag}\left(\exp(-\beta\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_\beta))\right)\left(\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_\beta)-\mathbf{1}_r\right) \xrightarrow[n\to\infty]{p} E\left[\exp\{-\beta g_i(\boldsymbol{\theta})\}(g_i(\boldsymbol{\theta})-1)^2\right], \tag{49}$$

其中 $E\left[\exp\{-\beta g_i(\boldsymbol{\theta})\}(g_i(\boldsymbol{\theta})-1)^2\right]=\varphi(-\beta)-2\frac{\partial}{\partial\beta}\varphi(-\beta)+\frac{\partial^2}{\partial\beta^2}\varphi(-\beta)=\frac{2(2\beta^2+1)}{(2\beta+1)^{5/2}}$，其中 $\varphi(s)=(1-2s)^{-\frac{1}{2}}$是 $g_i(\boldsymbol{\theta})\sim\chi^2_1$的矩母函数，在误差项服从正态分布时，除其他取值外，对$s=-\beta<0$也成立。考虑到（49）意味着

$$\tfrac{1}{n}\boldsymbol{v}^T(\widetilde{\boldsymbol{\theta}}_\beta)\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_\beta) \xrightarrow[n\to\infty]{p} E\left[\exp\{-\beta g_i(\boldsymbol{\theta})\}(g_i(\boldsymbol{\theta})-1)^2\right]-\frac{\beta^2}{(\beta+1)^3},$$

根据推论10 和斯卢茨基定理，可得所需结果。 ∎

(48) 的计算可通过对最小二乘法作适当解释来完成，具体如下。它等于 n 乘以将 $\boldsymbol{v}(\widetilde{\boldsymbol{\theta}}_\beta)$ 作为响应变量、将 $\mathbb{Z}_n$ 作为解释变量进行线性回归所得的判定系数。因此，经典 Koenker 检验统计量仍保留了原有的直观解释。在推论 16 中，已采用与 Koenker (1981) 类似的方法，证明了其渐近分布在误差项服从正态分布时成立。鉴于经典 Koenker检验在观测值不服从正态分布时已在线性分位数回归框架下得到证明（见 Koenker，1982），现在将证明 (48) 的渐近分布在更一般的条件下成立。线性回归模型 (5) 中误差项 (11) 的正态性假设可放宽为任意具有有限方差的中心化分布，且标准化误差平方的矩母函数存在二阶导数。

7 示例：房价数据

在 `wooldridge R` 软件包中，数据框 `hprice1` 与计量经济学中一个由 Wooldridge (2020) 提出的著名示例有关，该示例称为房价数据。该数据集包含 88 名个体的观测值和 10 个变量，代表波士顿地区售出房屋的特征及销售信息；这些数据采集自一个于 1990 年发布的公开数据库。我们考虑异方差线性模型（5），$\texttt{price}=\beta_0+\beta_1\,\texttt{bdrms}+\beta_2\,\texttt{lotsize}+\beta_3\,\texttt{sqrft}$，其中 `price` 为房价，单位为千美元（$y$），`bdrms`为卧室数量（$x_1$），`lotsize`为地块面积，单位为平方英尺（$x_2$），`sqrft`为房屋面积，单位为平方英尺（$x_3$）。图3 显示了存在

残差值最大的三个离群值，即数据框中标识编号为 43、72 和 76 的观测值。

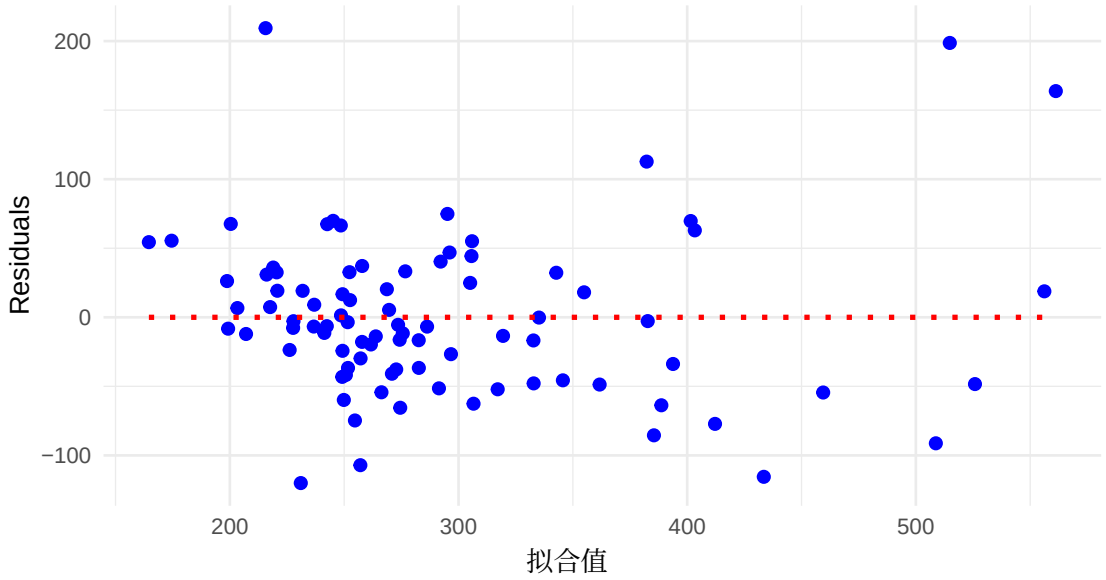


图 3：使用普通最小二乘法或最大似然估计进行估计时，房价数据的残差图（$\beta = 0$）。

针对完整数据或删除了前文提及的三个观测值后的数据，对调优参数区间为 $\beta \in [0, 0.75]$ 时 Breusch-Pagan 和 Koenker 的 β- 分数 LM $\tilde{}$ $\tilde{}$ 检验统计量 R_n、$_\beta$（$\boldsymbol{\theta}_\beta$）以及 Q_n、$_\beta$（$\boldsymbol{\theta}_\beta$）的表现进行了广泛研究。我们将重点关注两种不同来源的异方差，这取决于 $\check{\boldsymbol{z}}_i = (1, \boldsymbol{z}_i)^T$、$i = 1$、……、$n$ 的选择。情形 $\boldsymbol{x}_i = \check{\boldsymbol{z}}_i$、$i = 1$、……、$n$（$r = p = 3$）对应图 4 及表 1-2 左侧的研究情形；而图 5 及表 1-2 右侧则考虑了 White 加性异方差模型，

$$h(\boldsymbol{z}_i^T \boldsymbol{\alpha}) = 1 + \boldsymbol{z}_i^T \boldsymbol{\alpha} = 1 + \sum_{j=1}^{3} x_{ij}\alpha_j + \sum_{j=1}^{3}\sum_{k \geq j}^{3} x_{ij}x_{ik}\alpha_{jk},$$

其中 $\boldsymbol{z}_i = (\boldsymbol{x}_i^T, \ \mathrm{vech}^T(\boldsymbol{x}_i\boldsymbol{x}_i^T))^T$，$vech(\boldsymbol{x}_i\boldsymbol{x}_i)$ 是矩阵 $\boldsymbol{x}_i\boldsymbol{x}_i^T$的半向量化，且$r = p + {}^{p}{}_{(p+1)}$ $_2$ $= 9 > p = 3$。附录最后一节给出了若干计算细节，这些细节对于估计和检验统计量都很有用。根据 χ^2_r在 r 取 3 或 9（具体取值取决于异方差来源的选择）时的 95% 分位数，图 4-5中的虚线表示检验的拒绝阈值，该阈值基于渐近分布。

本示例的动机在于，人们怀疑经典检验（采用 $\beta = 0$ 的检验）可能会导致不恰当的决策。对于所考察的两种经典检验，同方差性原假设显然被拒绝；但删除离群值后，该假设则被接受。我们的目的是通过研究新提出的检验统计量在整个调优参数区间 $[0, 0.75]$ 上的表现，为所要作出的决策提供进一步的见解。$\widetilde{R}_{n,\beta}(\boldsymbol{\theta}_\beta)$ 和 $Q_{n,\beta}(\boldsymbol{\theta}_\beta)$ 在 $\beta > 0$ 时的优势在于，它们在存在离群值时具有抑制作用，从而避免删除离群值。不过，出于说明目的，分析并比较删除离群值时检验统计量的表现也很有意思。

鉴于图4 以及表格1-2左侧所示结果，我们有充分证据在 0.05 显著性水平下拒绝同方差性原假设；但从图 5和表格 1-2 右侧来看，由于 Koenker 的两个检验统计量并不一致，因此结论不甚明确：当 β-分数 LM 检验统计量在 $\beta \in (0.36, 0.53)$ 时，它处于接受域。然而，若将显著性水平设为 0.056，则在整个调优参数区间内，两个检验统计量都支持作出拒绝原假设的相同决策 $[0.0.75]$。这一决策与 Berenguer-Rico 和 Wilms（2021）所作的决策并不一致。事实上，他们的提议背后的方法也有所不同，因为该方法的构建以从数据中剔除离群值为优先考虑。

	情形 $\mathbb{X}_n = \breve{\mathbb{Z}}_n$			White 版本的 $\breve{\mathbb{Z}}_n$		
	$\beta = 0$	$\beta = 0.3$	$\beta = 0.6$	$\beta = 0$	$\beta = 0.3$	$\beta = 0.6$
$R_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)$	1.364e-06	3.501e-03	4.6799e-04	6.559e-12	1.179e-02	0.0128
$Q_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)$	2.782e-03	1.526e-02	6.9919e-03	9.952e-05	1.684e-02	0.0137

表 1：使用完整房价数据时，Breusch-Pagan 和 Koenker 的 β-分数 LM 检验统计量的 p 值。

	情形 $\mathbb{X}_n = \breve{\mathbb{Z}}_n$			White 版本的 $\breve{\mathbb{Z}}_n$		
	$\beta = 0$	$\beta = 0.3$	$\beta = 0.6$	$\beta = 0$	$\beta = 0.3$	$\beta = 0.6$
$R_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)$	0.0615	0.0252	0.0041	0.1743	0.0550	0.0618
$Q_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)$	0.0898	0.0275	0.0138	0.2576	0.0582	0.0548

表 2：删除离群值后的房价数据中，Breusch-Pagan 和 Koenker 的 β-分数 LM 检验统计量的 p 值。

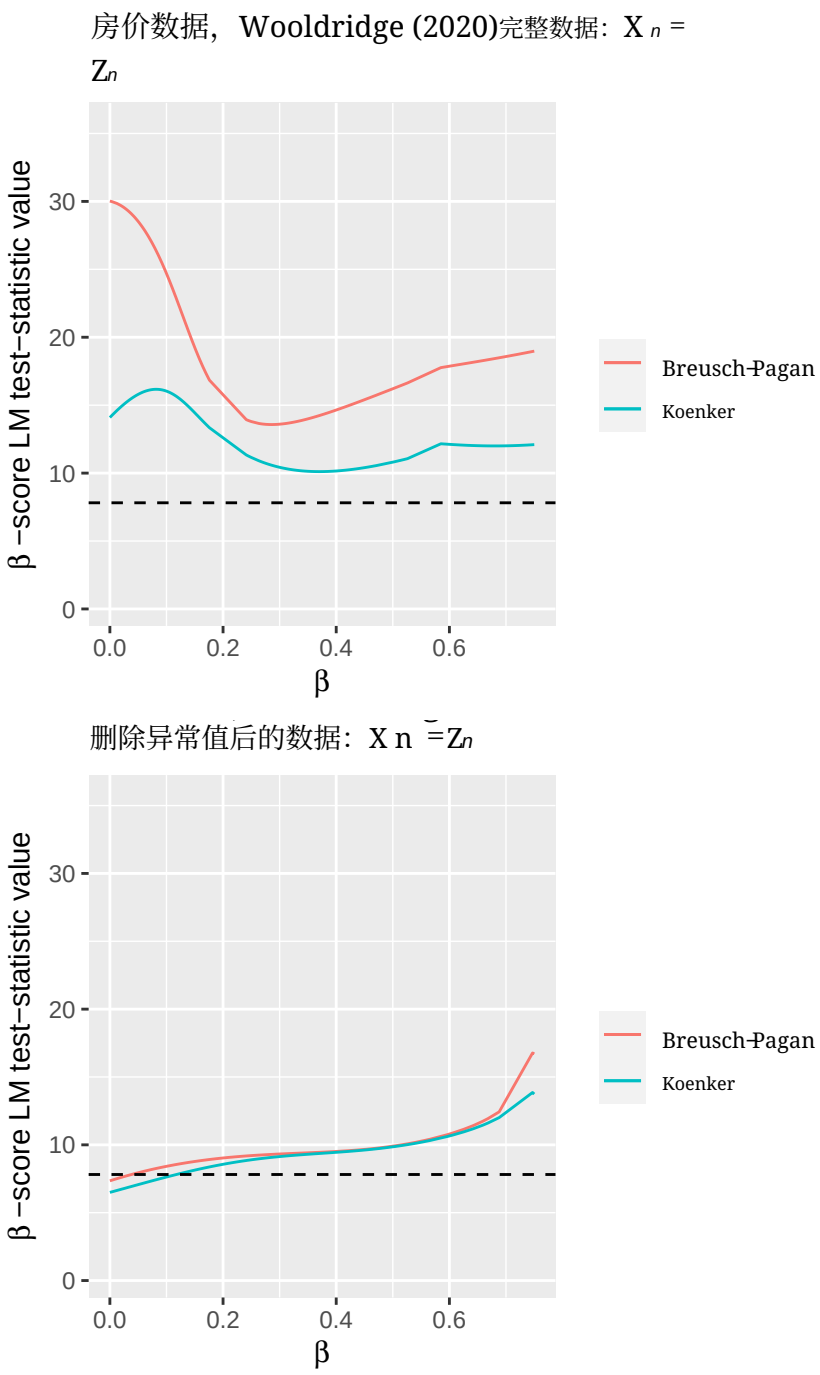


图 4：房价数据和情形 $\mathbb{X}_n = \mathbb{Z}_n$ 的 Breusch-Pagan 和 Koenker 的 β-分数 LM 检验统计量˘。

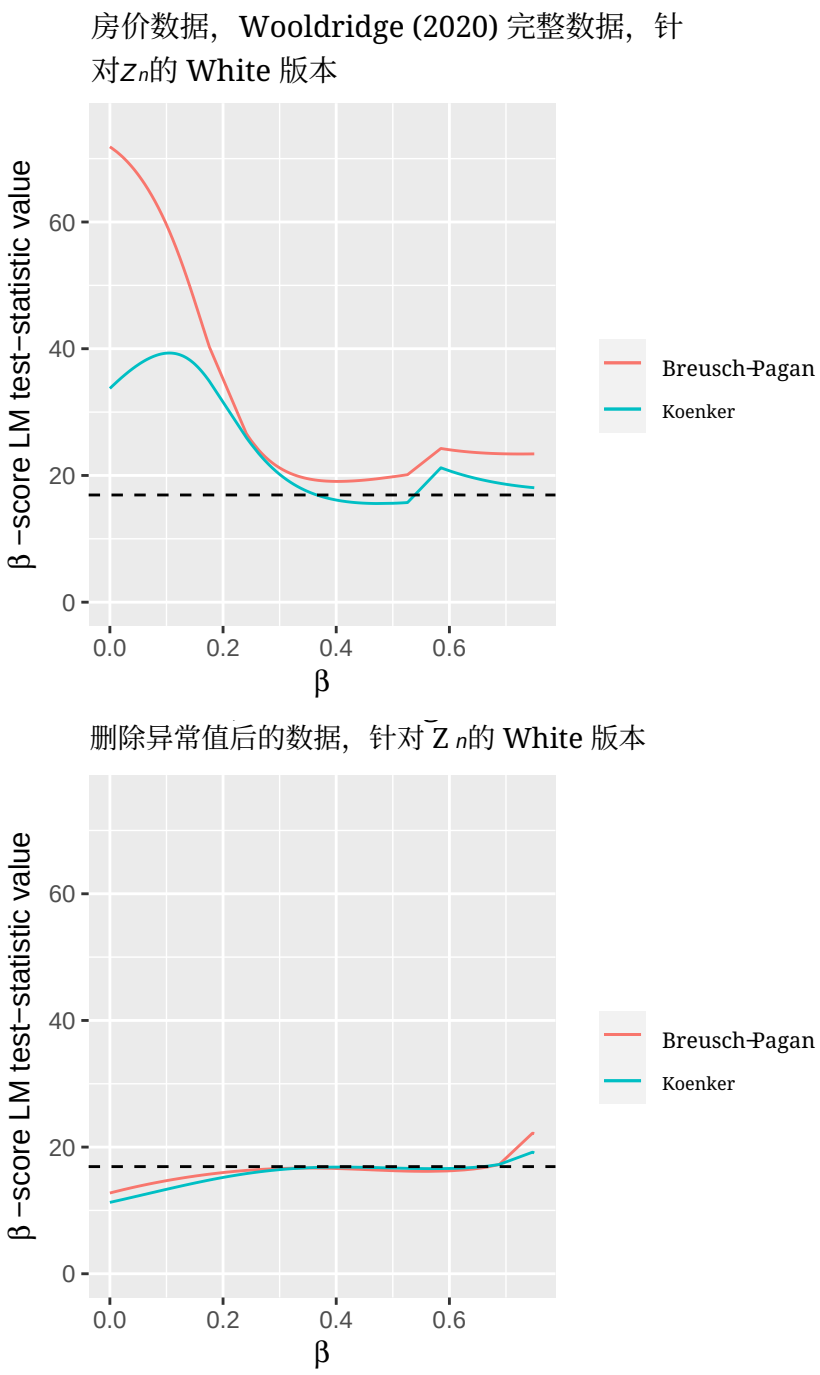


图 5：房价数据的 Breusch-Pagan 和 Koenker 的 β-分数 LM 检验统计量$\check{}$以及 $\mathbb{Z}_n$ 的 White 版本。

8 总结

本文提出了用于检验线性回归模型中异质性的 Breusch-Pagan β-分数 LM 检验。在提出该检验之前，我们基于独立但非同分布观测值的复合原假设框架，建立了推导该检验所需的一般理论，因为 Basu 等人（2022）仅考虑了同分布观测值。Breusch-Pagan β-分数 LM 检验统计量取决于调节参数 $\beta \geq 0$ ，而情形 $\beta = 0$ 则是情形 $\beta > 0$中 $\beta \to 0^+$ 的极限。该检验统计量族将经典 Breusch-Pagan LM 检验（Breusch 和 Pagan，1979）视为一个特例，即情形 $\beta = 0$。通过调节参数 $\beta \geq 0$，可以在效率和稳健性之间取得平衡： $\beta = 0$ 的效率最高，适用于没有离群值的干净数据；增大 $\beta > 0$的值则可提高检验统计量的稳健性，适用于存在离群值的情况。适当选择调节参数的值，可以在效率和稳健性之间取得折中。

自 Koenker (1981) 发表了一个针对正态性假设具有稳健性的修正版以来，该版本一直是稳健性的参考标准；而本文填补了提出一种对离群值具有稳健性的特定检验这一空白。此外，Koenker 的 β-分数检验也解决了这一挑战：如何推广我们提出的用于检验异质性的 β-分数 LM 检验，使其能够同时抵御正态性假设和离群值的影响。

当数据中存在离群值时，为避免采用基于 Breusch（1978）和 Godfrey（1978b）的经典 Breusch-Godfried LM 检验的有偏且低效版本得出结果，我们目前正在撰写一篇论文，拟提出用于检验线性回归模型误差项自相关的 β-分数 LM 检验。除了上述尚待解决的问题之外，基于 Breusch 和 Pagan（1980）的思路，本文也可成为后续具有挑战性的研究的关键依据和参考，为非同分布个体的 β-分数 LM 检验在计量经济学中的新应用提供借鉴。对此，Halunga 等（2017）近期提出了一种针对异方差具有稳健性的布鲁施-帕甘检验，用于动态面板数据中零截面相关的原假设。尽管他们的模型以及稳健技术都与我们的提议不同，但他们指出：“如果检验未能拒绝原假设，那么可以更有把握地得出结论：拒绝结果并非由异方差所致，此时宜采用普通最小二乘法估计。若拒绝原假设，则应采用适当的估计程序。”这一观点与本文模型的情境十分契合，而 β-分数 LM 检验可能是一种颇具吸引力的替代方法。

参考文献

[1] Alih, E. 和 Ong, H.C.（2015）。线性模型中的抗离群值异方差检验。应用统计学杂志，**42**，1617–1634。

[2] Aitchison, J. 和 Silvey, D.S.（1958）。受约束参数的极大似然估计。《数学统计年刊》，29，813–828。[3] Athreya, K.B. 和 Lahiri, S.N.（2006）。测度论与概率论。施普林格出版社。[4] Basu, A.、Ghosh, A.、Martin, N. 和 Pardo, L.（2022）。Rao检验的稳健推广。《商业与经济统计学期刊》，40，868–879。[5] Basu, A、Harris, I.R.、Hjort, N.L. 和 Jones, M.C.（1998）。通过最小化密度幂散度进行稳健且高效的估计。《生物计量学》，85，549–559。[6] Basu, A.、Mandal, A.、Martin, N. 和 Pardo, L.（2017）。基于密度幂散度的复合假设检验。Sankhya B，80，222–262。[7] Basu, A.、Shioya, H. 和 Park C.（2011）。统计推断：最小距离方法。博卡拉顿，CRC出版社[8] Beran, R. J.（1977）。参数模型的最小海灵格距离估计。《统计年刊》，5，445–463。´ Berenguer-Rico, V. 和 Wilms, I.（2021）。离群值剔除后的异方差检验。《计量经济学评论》，40，51–85。[9] Bickel, P.J.（1978）。稳健地使用残差 I：异方差与非线性检验。《统计年刊》，6，266–291。[10]Breusch, T. S.（1978）。动态线性模型中的自相关检验。Australian EconomicPapers，17，334–355。[11] Breusch, T.S. 和 Pagan, A.R.（1979）。异方差与随机系数变异的简单检验。《计量经济学》，47，1287–1294。[12] Breusch, T.S. 和 Pagan, A.R.（1980）。拉格朗日乘数检验及其在计量经济学模型设定中的应用。《经济研究评论》，47，239–253。[13] Boos, D. D.（1992）。关于广义得分检验。《美国统计学家》，46，327–333。[14] Carrol, R.J. 和 Ruppert, D.（1988）。回归中的变换与加权。CRC出版社。[15] Cook, R. D. 和 Weisberg, S.（1983）。回归中的异方差诊断。《生物计量学》，70，1–10。

[17] Cressie, N. 和 Read, T.R.C.（1984）。多项式拟合优度检验。《皇家统计学会杂志》*B* 辑，**46**，440–464。[18] Csisz´ar, I.（1967）。概率分布差异的信息型度量与间接观测。匈牙利数学科学研究，**2**，299–318。[19] Ghosh, A. 和 Basu, A.（2013）。使用密度幂散度对独立非同质观测值进行稳健估计及其在线性回归中的应用。电子统计学期刊，**7**，2420–2456。[20] Ghosh, A.、Mandal, A.、Mart´ín, N. 和 Pardo, L.（2016）。稳健瓦尔德型检验的影响分析。多元分析杂志，**147**，102–126。[21] Godfrey, L.G.（1978a）。乘法异方差检验。《计量经济学杂志》，**8**，227-236。[22] Godfrey, L.G.（1978b）。当回归变量包含滞后因变量时，针对一般自回归和移动平均误差模型的检验。《计量经济学》，**46**，1293–1301。[23] Godfrey, L. G.（1989）。《计量经济学中的设定错误检验：拉格朗日乘数原理与其他方法》。剑桥大学出版社，纽约。[24] Halunga, A.G.、Orme, C.D、Yamagata, T.（2017）。动态面板数据模型中同期相关的异方差稳健 Breusch-Pagan 检验。《计量经济学杂志》。198，209–230。[25] Hampel, F.R.（1968）。对稳健估计理论的贡献。博士论文，加州大学伯克利分校。[26] Hampel, F.R.（1974）。影响曲线及其在稳健估计中的作用。《美国统计学会杂志》，**69**，383–393。[27] Hampel, F.R.、E.M. Ronchetti、P.J. Rousseeuw 和 W.A. Stahel（1986）。《稳健统计：基于影响函数的方法》。Wiley，纽约。[28] Hannan, E.J.（1956）。基于多重相关的若干检验的渐近功效。《皇家统计学会杂志》*B*辑，**18**，227–233。[29] Honda, Y.（1988）。异方差拉格朗日乘数检验的检验规模校正。《计量经济学杂志》，**38**，375–386。[30] Kalina, J.（2011）稳健回归中的异方差检验。《经济、商业与信息通信技术研究期刊，**4**，25–28。

[31] Kim, B. 和 Lee, S.（2013）。SCOMDY模型中Copula参数的稳健估计。《时间序列分析杂志》，34，302–314。[32] Kim, B. 和 Lee, S.（2017）。基于密度幂散度的零膨胀泊松自回归模型稳健估计。《统计计算与模拟杂志》，87，2981–2996。[33] Kim, B. 和 Lee, S.（2020）。一般整数值时间序列模型的稳健估计。《统计数学研究所年刊》，72(6)，1371-1396。[34] Kim, B.（2018）。基于最小密度幂散度估计量的GARCH模型稳健最大熵检验。《经济学快报》，162，93–97。[35] Lee, S. 和 Song, J.（2009）。GARCH模型的最小密度幂散度估计量。《TEST》18，316–341。[36] Koenker, R.（1981）。关于异方差检验学生化的一则说明。《计量经济学杂志》，17，107–112。[37] Koenker, R. 和 Bassett, G.（1982）。基于回归分位数的异方差稳健检验。《计量经济学》，50，43–61。[38] Koenker, R. 和 Bassett, G.（1982）。线性假设检验与 l_1 估计。《计量经济学》，50，1577–1583。[39] Krasker, W.S. 和 Welsch, R.E.（1982）。高效的有界影响回归估计。《美国统计学会杂志》，77，595–604。[40] Kolmogorov, A. N.（1933）。关于分布律的经验确定。《意大利精算师学会杂志》，4，83-91。[41] Kmenta, J.（1986）。《计量经济学原理》。麦克米伦。[42] Lyon, J. D. 和 Tsai, C.L.（1996）。异方差检验的比较。《皇家统计学会杂志》D辑（《统计学家》），45，337–349。[43] Lu, T. T. 和 Shiou, S. H.（2002）。 2×2 分块矩阵的逆。计算与数学应用，43，119–129。[44] Mart´ın, N.（2021）。相关矩阵上的Rao得分检验。arXiv，统计理论（math.ST），https://doi.org/10.48550/arXiv.2012.14238。[45] Mahalanobis, P. C.（1936）。统计学中的广义距离。《印度国家科学研究所会刊》，2，49–55。

[46] Markatou, M., Karlis, D., 和 Ding, Y.（2021）。《基于距离的统计推断》。《统计学及其应用年度评论》，**8**，301–327。[47] A. Nuttall（1975）。一些涉及 Q_M 函数的积分。《*IEEE* 信息论汇刊》，**21**，95–96。[48]Pardo, L.（2006）。基于散度度量的统计推断。$Chapman\ \ andHall/CRC$。[49] Rao, C. R., Toutenburg, H., Shalabh 和 Heumann, C.（2008）。线性模型及其推广。最小二乘法及其替代方法。施普林格。[50] Salibian-Barrera, M., Van Aelst, S. 和 Yohai, V.J.（2016）。基于 τ 估计量的线性回归模型稳健检验。《计算统计学与数据分析》，**93**，436–455。[51] Silvey, S.D.（1959）。拉格朗日乘数检验。《数学统计年刊》，**30**，389–407。[52] Waldman, D.M.（1983）。关于 White 检验与 Godfrey/Breusch-Pagan 异方差检验的一种变体之间代数等价性的注记。《经济学快报》，13，197–200。[53] White. H.（1980）。异方差稳健协方差矩阵估计量及异方差直接检验。《计量经济学》，48，817-838。

附录

定理4的证明

我们将遵循 Davidson 和 MacKinnon（2021，第275页）在第8.9节中阐述的类似方案，该节专门讨论经典 LM 检验，但将其调整为适用于独立且非同分布的个体以及基于 DPD 的估计函数。我们可以从考虑 $\sim \boldsymbol{u}_{i,\beta}(Y_i;\ \boldsymbol{\theta}_\beta)$ 在点 $\boldsymbol{\theta}_0$ for $i = 1, \ldots, n$ 处的泰勒展开开始。

$$\boldsymbol{u}_{i,\beta}(Y_i;\widetilde{\boldsymbol{\theta}}_\beta) = \boldsymbol{u}_{i,\beta}(Y_i;\boldsymbol{\theta}_0) + \frac{\partial}{\partial \boldsymbol{\theta}}\ \boldsymbol{u}_{i,\boldsymbol{\beta}}^T(Y_i;\boldsymbol{\theta})\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}(\widetilde{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_0) + o(\left\|\widetilde{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_0\right\|^2 \mathbf{1}_r),$$

我们得到

$$\sqrt{n}\boldsymbol{u}_{i,\beta}(Y_i;\widetilde{\boldsymbol{\theta}}_\beta) = \sqrt{n}\boldsymbol{u}_{i,\beta}(Y_i;\boldsymbol{\theta}_0) + \sqrt{n}\frac{\partial}{\partial \boldsymbol{\theta}}\ \boldsymbol{u}_{i,\beta}^T(Y_i;\boldsymbol{\theta})\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}(\widetilde{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_0) + o(\sqrt{n}\left\|\widetilde{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_0\right\|^2 \mathbf{1}_r),$$

由此可得

$$\sqrt{n}\boldsymbol{U}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)=\sqrt{n}\boldsymbol{U}_{n,\beta}(\boldsymbol{\theta}_0)+\frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{U}^T_{n,\beta}(\boldsymbol{\theta}_0)\sqrt{n}(\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_0)+o(\sqrt{n}\left\|\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_0\right\|^2\mathbf{1}_r).$$

根据（2）调整左侧表达式，并考虑马尔可夫弱大数定律

$$\frac{\partial}{\partial\boldsymbol{\theta}}\boldsymbol{U}^T_{n,\beta}(\boldsymbol{\theta}_0)\xrightarrow[n\to\infty]{p}-\boldsymbol{J}_\beta(\boldsymbol{\theta}_0),$$

可得

$$\sqrt{n}\boldsymbol{M}(\widetilde{\boldsymbol{\theta}}_\beta)\boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)=\sqrt{n}\boldsymbol{U}_{n,\beta}(\boldsymbol{\theta}_0)-\boldsymbol{J}_\beta(\boldsymbol{\theta}_0)\sqrt{n}(\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_0)+o(\sqrt{n}\left\|\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_0\right\|^2\mathbf{1}_r).\tag{50}$$

~另一方面，根据 $\boldsymbol{m}(\boldsymbol{\theta}_\beta)$ 在点 $\boldsymbol{\theta}_0$ 处的泰勒展开，

$$\boldsymbol{m}(\widetilde{\boldsymbol{\theta}}_\beta)=\boldsymbol{m}(\boldsymbol{\theta}_0)+\boldsymbol{M}^T(\boldsymbol{\theta}_0)(\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_0)+o(\left\|\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_0\right\|^2\mathbf{1}_r),$$

并考虑到（3），可得

$$\mathbf{0}_r=\boldsymbol{M}^T(\boldsymbol{\theta}_0)\sqrt{n}(\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_0)+o(\sqrt{n}\left\|\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_0\right\|^2\mathbf{1}_r).\tag{51}$$

由方程（50）-（51）可得

$$\begin{bmatrix}\boldsymbol{J}_\beta(\boldsymbol{\theta}_0) & \boldsymbol{M}(\boldsymbol{\theta}_0)\\ \boldsymbol{M}^T(\boldsymbol{\theta}_0) & \mathbf{0}_{p\times p}\end{bmatrix}\sqrt{n}\begin{bmatrix}\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_0\\ \boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\end{bmatrix}=\begin{bmatrix}\sqrt{n}\boldsymbol{U}_{n,\beta}(\boldsymbol{\theta}_0)\\ \mathbf{0}_r\end{bmatrix}+o(\sqrt{n}\left\|\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_0\right\|^2\mathbf{1}_r).$$

因此，

$$\begin{aligned}\sqrt{n}\begin{bmatrix}\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_0\\ \boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\end{bmatrix}&=\begin{bmatrix}\boldsymbol{J}_\beta(\boldsymbol{\theta}_0) & \boldsymbol{M}(\boldsymbol{\theta}_0)\\ \boldsymbol{M}^T(\boldsymbol{\theta}_0) & \mathbf{0}_{r\times r}\end{bmatrix}^{-1}\begin{bmatrix}\sqrt{n}\boldsymbol{U}_{n,\beta}(\boldsymbol{\theta}_0)\\ \mathbf{0}_r\end{bmatrix}+o(\sqrt{n}\left\|\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_0\right\|^2\mathbf{1}_r)\\&=\begin{bmatrix}\boldsymbol{\Sigma}_{11,\beta}(\boldsymbol{\theta}_0)\sqrt{n}\boldsymbol{U}_{n,\beta}(\boldsymbol{\theta}_0)\\ \boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0)\sqrt{n}\boldsymbol{U}_{n,\beta}(\boldsymbol{\theta}_0)\end{bmatrix}+o(\sqrt{n}\left\|\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_0\right\|^2\mathbf{1}_r),\end{aligned}\tag{52}$$

with

$$\begin{bmatrix}\boldsymbol{J}_\beta(\boldsymbol{\theta}_0) & \boldsymbol{M}(\boldsymbol{\theta}_0)\\ \boldsymbol{M}^T(\boldsymbol{\theta}_0) & \mathbf{0}_{r\times r}\end{bmatrix}^{-1}=\begin{bmatrix}\boldsymbol{\Sigma}_{11,\beta}(\boldsymbol{\theta}_0) & \boldsymbol{\Sigma}^T_{21,\beta}(\boldsymbol{\theta}_0)\\ \boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0) & \boldsymbol{\Sigma}_{22,\beta}(\boldsymbol{\theta}_0)\end{bmatrix},$$

其中

$$\begin{aligned}\boldsymbol{\Sigma}_{11,\beta}(\boldsymbol{\theta}_0)&=\boldsymbol{J}^{-1}_\beta(\boldsymbol{\theta}_0)+\boldsymbol{J}^{-1}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{22,\beta}(\boldsymbol{\theta}_0)\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{J}^{-1}_\beta(\boldsymbol{\theta}_0)\\&=\boldsymbol{J}^{-1}_\beta(\boldsymbol{\theta}_0)-\boldsymbol{J}^{-1}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{J}^{-1}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0))^{-1}\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{J}^{-1}_\beta(\boldsymbol{\theta}_0),\\\boldsymbol{\Sigma}_{22,\beta}(\boldsymbol{\theta}_0)&=-(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{J}^{-1}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0))^{-1},\\\boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0)&=-\boldsymbol{\Sigma}_{22,\beta}(\boldsymbol{\theta}_0)\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{J}^{-1}_\beta(\boldsymbol{\theta}_0)\\&=(\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{J}^{-1}_\beta(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0))^{-1}\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{J}^{-1}_\beta(\boldsymbol{\theta}_0),\end{aligned}$$

并且在（18）-（19）的存在性假设下，还验证了适用于独立非同分布观测的林德伯格中心极限定理的多元版本，

$$\sqrt{n}\boldsymbol{U}_{n,\beta}(\boldsymbol{\theta}_0) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(\mathbf{0}_p, \boldsymbol{K}_\beta(\boldsymbol{\theta})). \tag{53}$$

鉴于定义3中给出的关于$\boldsymbol{K}_{n,\beta}(\boldsymbol{\theta})$的表达式，这只是中心化样本的方差估计量，其依据如下

$$\begin{aligned}
\bar{\boldsymbol{K}}_{n,\beta}(\boldsymbol{\theta}) &= n\mathrm{E}\left[\boldsymbol{U}_{n,\beta}(\boldsymbol{\theta})\boldsymbol{U}_{n,\beta}^T(\boldsymbol{\theta})\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n}\mathrm{E}[\boldsymbol{u}_{i,\beta}(Y_i;\boldsymbol{\theta})]\mathrm{E}[\boldsymbol{u}_{j,\boldsymbol{\theta}}^T(Y_i;\boldsymbol{\theta})] + \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}[\boldsymbol{u}_{i,\beta}(Y_i;\boldsymbol{\theta})\boldsymbol{u}_{i,\beta}^T(Y_i;\boldsymbol{\theta})] \\
&= \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{K}_{i,\beta}(\boldsymbol{\theta}),
\end{aligned}$$

其中 $\mathrm{E}[\boldsymbol{u}_{i,\beta}(Y_i;\boldsymbol{\theta})] = \mathbf{0}_p, i = 1, \ldots\ldots, n$。表达式 $\bar{\boldsymbol{K}}_{n,\beta}(\boldsymbol{\theta}) = \bar{\boldsymbol{J}}_{n,2\beta}(\boldsymbol{\theta}) - \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\xi}_{i,\beta}(\boldsymbol{\theta})\boldsymbol{\xi}_{i,\beta}^T(\boldsymbol{\theta})$，可直接由.....得到

$$\boldsymbol{K}_{i,\beta}(\boldsymbol{\theta}) = \boldsymbol{J}_{i,2\beta}(\boldsymbol{\theta}) - \boldsymbol{\xi}_{i,\beta}(\boldsymbol{\theta})\boldsymbol{\xi}_{i,\beta}^T(\boldsymbol{\theta}), \quad i = 1, \ldots, n,$$

Basu等（1998）针对同分布观测值建立了这一结果。最后，由（52）-（53）可得所需结果。

B 定理5的证明

由（53）可得

$$\begin{aligned}
&\sqrt{n}\boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}\left(\mathbf{0}_r, \boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0)\boldsymbol{K}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{21,\beta}^T(\boldsymbol{\theta}_0)\right), \\
&\left(\bar{\boldsymbol{\Sigma}}_{n,21,\beta}^T(\widetilde{\boldsymbol{\theta}}_\beta)\bar{\boldsymbol{K}}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\bar{\boldsymbol{\Sigma}}_{n,21,\beta}^T(\widetilde{\boldsymbol{\theta}}_\beta)\right)^{-\frac{1}{2}}\sqrt{n}\boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}\left(\mathbf{0}_r, \boldsymbol{I}_r\right), \\
&R_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta) = n\boldsymbol{\lambda}_{n,\beta}^T(\widetilde{\boldsymbol{\theta}}_\beta)\left(\bar{\boldsymbol{\Sigma}}_{n,22,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\bar{\boldsymbol{K}}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\bar{\boldsymbol{\Sigma}}_{n,22,\beta}^T(\widetilde{\boldsymbol{\theta}}_\beta)\right)^{-1}\boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta) \xrightarrow[n\to\infty]{\mathcal{L}} \chi_r^2.
\end{aligned}$$

C 定理6的证明

令 $\boldsymbol{\theta}_n$为方程 $\sqrt{n}\boldsymbol{m}(\boldsymbol{\theta}_n) = \boldsymbol{\delta}$ 对任意 $n \in \mathbb{N}$的解。按照定理 A 证明中的相同思路，但将 $\boldsymbol{\theta}_0$ 替换为 $\boldsymbol{\theta}_n$，可得

$$\boldsymbol{m}(\widetilde{\boldsymbol{\theta}}_\beta) = \boldsymbol{m}(\boldsymbol{\theta}_n) + \boldsymbol{M}^T(\boldsymbol{\theta}_0)(\widetilde{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_n) + o(\left\|\widetilde{\boldsymbol{\theta}}_\beta - \boldsymbol{\theta}_n\right\|^2 \mathbf{1}_r),$$

因此

$$\sqrt{n}\boldsymbol{m}(\widetilde{\boldsymbol{\theta}}_\beta)=\sqrt{n}\boldsymbol{m}(\boldsymbol{\theta}_n)+\boldsymbol{M}^T(\boldsymbol{\theta}_0)\sqrt{n}(\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_n)+o(\sqrt{n}\left\|\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_n\right\|^2\mathbf{1}_r), \tag{54}$$

其中

$$\sqrt{n}\boldsymbol{m}(\widetilde{\boldsymbol{\theta}}_\beta)=\sqrt{n}\boldsymbol{m}(\boldsymbol{\theta}_0)+o_p(\mathbf{1}_r),$$

然后

$$\begin{aligned}\sqrt{n}\begin{bmatrix}\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_n\\ \boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\end{bmatrix}&=\begin{bmatrix}\boldsymbol{J}_\beta(\boldsymbol{\theta}_0) & \boldsymbol{M}(\boldsymbol{\theta}_0)\\ \boldsymbol{M}^T(\boldsymbol{\theta}_0) & \mathbf{0}_{r\times r}\end{bmatrix}^{-1}\begin{bmatrix}\sqrt{n}\boldsymbol{U}_{n,\beta}(\boldsymbol{\theta}_0)\\ -\boldsymbol{\delta}\end{bmatrix}+o(\sqrt{n}\left\|\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_n\right\|^2\mathbf{1}_r)\\ \sqrt{n}\boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)&=\boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0)\sqrt{n}\boldsymbol{U}_{n,\beta}(\boldsymbol{\theta}_0)-\boldsymbol{\Sigma}_{22,\beta}(\boldsymbol{\theta}_0)\boldsymbol{\delta}+o(\sqrt{n}\left\|\widetilde{\boldsymbol{\theta}}_\beta-\boldsymbol{\theta}_n\right\|^2\mathbf{1}_r).\end{aligned} \tag{55}$$

最后，由（53）可得

$$\begin{aligned}&\sqrt{n}\boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\xrightarrow[n\to\infty]{\mathcal{L}}\mathcal{N}\left(-\boldsymbol{\Sigma}_{22,\beta}(\boldsymbol{\theta}_0)\boldsymbol{\delta},\boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0)\boldsymbol{K}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{21,\beta}^T(\boldsymbol{\theta}_0)\right),\\ &\left(\bar{\boldsymbol{\Sigma}}_{n,21,\beta}^T(\widetilde{\boldsymbol{\theta}}_\beta)\bar{\boldsymbol{K}}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\bar{\boldsymbol{\Sigma}}_{n,21,\beta}^T(\widetilde{\boldsymbol{\theta}}_\beta)\right)^{-\frac{1}{2}}\sqrt{n}\boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\xrightarrow[n\to\infty]{\mathcal{L}}\mathcal{N}\left(\boldsymbol{\mu}_\beta(\boldsymbol{\delta}),\boldsymbol{I}_r\right),\\ &R_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)=n\boldsymbol{\lambda}_{n,\beta}^T(\widetilde{\boldsymbol{\theta}}_\beta)\left(\bar{\boldsymbol{\Sigma}}_{n,22,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\bar{\boldsymbol{K}}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\bar{\boldsymbol{\Sigma}}_{n,22,\beta}^T(\widetilde{\boldsymbol{\theta}}_\beta)\right)^{-1}\boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\xrightarrow[n\to\infty]{\mathcal{L}}\chi_r^2(\nu_\beta(\boldsymbol{\theta}_0,\boldsymbol{\delta})),\end{aligned}$$

with

$$\begin{aligned}\boldsymbol{\mu}_\beta(\boldsymbol{\delta})&=-\left(\boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0)\boldsymbol{K}(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{21,\beta}^T(\boldsymbol{\theta}_0)\right)^{-\frac{1}{2}}\boldsymbol{\Sigma}_{22,\beta}(\boldsymbol{\theta}_0)\boldsymbol{\delta}=\boldsymbol{E}_\beta^{\frac{1}{2}}(\boldsymbol{\theta})\boldsymbol{\delta},\\ \nu_\beta(\boldsymbol{\theta},\boldsymbol{\delta})&=\boldsymbol{\delta}^T\boldsymbol{E}_\beta\boldsymbol{\delta},\end{aligned}$$

and

$$\boldsymbol{E}_\beta(\boldsymbol{\theta})=\left(\boldsymbol{M}^T(\boldsymbol{\theta})\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta})\boldsymbol{K}(\boldsymbol{\theta})\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta})\boldsymbol{M}(\boldsymbol{\theta})\right)^{-1}$$

是检验效能。

定理11的证明

根据定理4，该条件成立

$$\sqrt{n}\boldsymbol{\lambda}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\xrightarrow[n\to\infty]{\mathcal{L}}\mathcal{N}\left(\mathbf{0}_p,\boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0)\boldsymbol{K}_\beta(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{21,\beta}^T(\boldsymbol{\theta}_0)\right),$$

with

$$\begin{aligned}\boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0)\boldsymbol{K}_\beta(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{21,\beta}^T(\boldsymbol{\theta}_0)&=\boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0)\mathrm{E}\left[\boldsymbol{U}_{n,\beta}\left(\boldsymbol{\theta}_0\right)\boldsymbol{U}_{n,\beta}^T\left(\boldsymbol{\theta}_0\right)\right]\boldsymbol{\Sigma}_{21,\beta}^T(\boldsymbol{\theta}_0)\\ &=\mathrm{E}\left[\mathcal{IF}\left(\boldsymbol{y},\boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\right)\mathcal{IF}_i^T\left(\boldsymbol{y},\boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\right)\right],\end{aligned}$$

$\boldsymbol{y} = (y_1, \ldots, y_n)^T$。因此，

与拉格朗日乘子对应的所有个体的影响函数由函数 $\mathbb{R}^n \longrightarrow \mathbb{R}$ 给出，使得

$$\mathcal{IF}\left(\boldsymbol{y}, \boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\right) = \frac{1}{n}\sum_{i=1}^{n}\mathcal{IF}_i\left(y_i, \boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\right),$$

其中

$$\mathcal{IF}_i\left(y_i, \boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\right) = \boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0)\boldsymbol{u}_{i,\beta}\left(y_i; \boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right), \quad i = 1, \ldots, n; \tag{56}$$

i.e.

$$\mathcal{IF}\left(\boldsymbol{y}, \boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\right) = \boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0)\boldsymbol{U}_{n,\beta}\left(\boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right).$$

$$\begin{aligned}
&\mathcal{IF}_i^{(2)}\left(y_i, R_{n,\beta}\left(\boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\right) = \\
&= \mathcal{IF}_i^T\left(y_i, \boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\right)\left[\boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0)\boldsymbol{K}_\beta(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{21,\beta}^T(\boldsymbol{\theta}_0)\right]^{-1}\mathcal{IF}_i\left(y_i, \boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\right) \\
&= \boldsymbol{u}_{i,\beta}^T\left(y_i; \boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\left[\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\boldsymbol{K}_\beta(\boldsymbol{\theta}_0)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\right]^{-1} \\
&\times \boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\boldsymbol{u}_{i,\beta}\left(y_i; \boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right) \\
&= \left\|\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\boldsymbol{u}_{i,\beta}\left(y_i; \boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\right\|^2_{\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\boldsymbol{K}_\beta(\boldsymbol{\theta}_0)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)}.
\end{aligned}$$

$$\begin{aligned}
&\mathcal{IF}_i^{(2)}\left(y_i, R_{n,\beta}\left(\boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\right) \\
&= \mathcal{IF}_i^T\left(y_i, \boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\right)\left[\boldsymbol{\Sigma}_{21,\beta}(\boldsymbol{\theta}_0)\boldsymbol{K}_\beta(\boldsymbol{\theta}_0)\boldsymbol{\Sigma}_{21,\beta}^T(\boldsymbol{\theta}_0)\right]^{-1}\mathcal{IF}_i\left(y_i, \boldsymbol{\lambda}_{n,\beta}\left(\boldsymbol{T}_{n,\beta}(\underline{F}_{\boldsymbol{\theta}_2})\right)\right) \\
&= \boldsymbol{d}_{i,\beta}^T(\boldsymbol{\theta}_0)\left[\boldsymbol{M}^T(\boldsymbol{\theta}_0)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\boldsymbol{K}_\beta(\boldsymbol{\theta}_0)\boldsymbol{J}_\beta^{-1}(\boldsymbol{\theta}_0)\boldsymbol{M}(\boldsymbol{\theta}_0)\right]^{-1}\boldsymbol{d}_{i,\beta}(\boldsymbol{\theta}_0),
\end{aligned}$$

E 获取 Breusch-Pagan 和 Koenker 的 β-分数检验统计量的计算注意事项

本节将介绍所采用的计算方法的若干细节，包括一些数学工具以及关于 R 代码的一些细节。

E.1 情形： $\beta = 0$

E.1.1 $\widetilde{\boldsymbol{\theta}} = (\widetilde{\boldsymbol{\beta}}, \widetilde{\sigma}^2)^T$、 $\widetilde{\widetilde{\boldsymbol{\theta}}} = (\widetilde{\boldsymbol{\beta}}, \widetilde{\widetilde{\sigma}}^2)^T$的最大似然估计量

考虑 $\mathbb{X}_n^T\mathbb{X}_n$ 的 Cholesky 分解，即 $\mathbb{X}_n^T\mathbb{X} = \mathbb{L}_n\mathbb{L}_n^T$其中 $\mathbb{L}_n$是一个下三角矩阵。考虑方程组

$$\begin{aligned} \mathbb{X}_n^T\mathbb{X}_n\widetilde{\boldsymbol{\beta}} &= \mathbb{X}_n^T\boldsymbol{Y}, \\ \mathbb{L}_n\mathbb{L}_n^T\widetilde{\boldsymbol{\beta}} &= \mathbb{X}_n^T\boldsymbol{Y}, \end{aligned} \tag{57}$$

分两步求解

$$\begin{cases} \mathbb{L}_n\widetilde{\widetilde{\boldsymbol{\beta}}} = \mathbb{X}_n^T\boldsymbol{Y}, \\ \mathbb{L}_n^T\widetilde{\boldsymbol{\beta}} = \widetilde{\widetilde{\boldsymbol{\beta}}}, \end{cases}$$

第一步为前代，第二步为回代。

E.1.2 经典 Breusch Pagan 和 Koenker检验统计量

考虑 $v_i(\widetilde{\boldsymbol{\theta}}) = g_i(\widetilde{\boldsymbol{\theta}}) - 1$, $g_i(\widetilde{\boldsymbol{\theta}}) = \frac{\varepsilon_i^2(\widetilde{\boldsymbol{\beta}})}{\widetilde{\sigma}^2}$, $i = 1, \ldots, n$ 以及 $\widetilde{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n \varepsilon_i^2(\widetilde{\boldsymbol{\beta}})$。我们必须推导一个方差分析:

$$\begin{aligned} \boldsymbol{v}(\widetilde{\boldsymbol{\theta}}) &= \breve{\mathbb{Z}}_n\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \\ &= \mathbf{1}_n\gamma_0 + \mathbb{Z}_n\boldsymbol{\alpha} + \boldsymbol{\varepsilon}, \end{aligned}$$

其中

$$\begin{aligned} \boldsymbol{v}(\widetilde{\boldsymbol{\theta}}) &= \boldsymbol{g}(\widetilde{\boldsymbol{\theta}}) - \mathbf{1}_n, \\ \boldsymbol{g}(\widetilde{\boldsymbol{\theta}}) &= \frac{\boldsymbol{\varepsilon}(\widetilde{\boldsymbol{\beta}}) \odot \boldsymbol{\varepsilon}(\widetilde{\boldsymbol{\beta}})}{\widetilde{\sigma}^2}, \\ \boldsymbol{\varepsilon}(\widetilde{\boldsymbol{\beta}}) &= \boldsymbol{Y} - \mathbb{X}_n\widetilde{\boldsymbol{\beta}}, \end{aligned}$$

$\odot$ 是逐元素 $\breve{}$是两个向量的逐元素乘积， $\widetilde{\sigma}^2$ 是 $\boldsymbol{v}^T(\boldsymbol{\theta})\mathbf{1}_n = 0$的解。然后

$$\begin{aligned} R_n(\widetilde{\boldsymbol{\theta}}) &= ESS(\widetilde{\boldsymbol{\theta}})/2, \\ Q_n(\widetilde{\boldsymbol{\theta}}) &= \frac{ESS(\widetilde{\boldsymbol{\theta}})}{TSS(\widetilde{\boldsymbol{\theta}})/n} = nR_{OLS}^2(\widetilde{\boldsymbol{\theta}}), \end{aligned}$$

分别是经典的 Breusch Pagan 检验统计量和 Koenker检验统计量，它们渐近服从具有 r 个自由度的卡方随机变量（满$\breve{}$秩矩阵 $\mathbb{Z}_n$ 的列数）。在非正态误差下， Q_n（$\boldsymbol{\theta}$）的相同渐近分布仍然$\breve{}$ $\breve{}$有效，但 TSS（$\boldsymbol{\theta}$）$/n$几乎必然不必趋于 2。 TSS（$\boldsymbol{\theta}$）$/n$的值是误差分布峰度减 1 的估计值。

E.2 情形： $\beta > 0$

E.2.1 $\widetilde{\boldsymbol{\theta}} = (\widetilde{\boldsymbol{\beta}}, \widetilde{\sigma}^2)^T$、$\widetilde{\boldsymbol{\theta}}_\beta = (\widetilde{\boldsymbol{\beta}}_\beta, \widetilde{\sigma}^2_\beta)^T$的最小 DPD 估计

其解为

$$\mathbb{X}_n^T \boldsymbol{D}_\beta(\widetilde{\boldsymbol{\theta}}_\beta)\mathbb{X}_n\widetilde{\boldsymbol{\beta}} = \mathbb{X}_n^T \boldsymbol{D}_\beta(\widetilde{\boldsymbol{\theta}}_\beta)\boldsymbol{Y}, \tag{58}$$

其中

$$\boldsymbol{D}_\beta(\widetilde{\boldsymbol{\theta}}_\beta) = \mathrm{diag}(\exp(-\tfrac{\beta}{2}\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_\beta))),$$

将会 递归求得。（58）将以与（57）相同的方式求解，取 g

$$\mathbb{X}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta) = \boldsymbol{D}_\beta^{1/2}(\widetilde{\boldsymbol{\theta}}_\beta)\mathbb{X}_n,$$
$$\boldsymbol{Y}(\widetilde{\boldsymbol{\theta}}_\beta) = \boldsymbol{D}_\beta^{1/2}(\widetilde{\boldsymbol{\theta}}_\beta)\boldsymbol{Y},$$

其中

$$\boldsymbol{D}_\beta^{1/2}(\widetilde{\boldsymbol{\theta}}_\beta) = \mathrm{diag}(\exp(-\tfrac{\beta}{4}\boldsymbol{g}(\widetilde{\boldsymbol{\theta}}_\beta))),$$

即（57）变为

$$\mathbb{X}_{n,\beta}^T(\widetilde{\boldsymbol{\theta}}_\beta)\mathbb{X}_{n,\beta}(\widetilde{\boldsymbol{\theta}}_\beta)\widetilde{\boldsymbol{\beta}} = \mathbb{X}_n^T(\widetilde{\boldsymbol{\theta}}_\beta)\boldsymbol{Y}(\widetilde{\boldsymbol{\theta}}_\beta).$$

在每次迭代中，σ^2 通过求解 $\boldsymbol{v}^T(\boldsymbol{\theta}_\beta)\mathbf{1}_n = 0$ 并将其表示为 σ^2的函数来更新，即 $f(\sigma^2) = 0$，其中

$$f(\sigma^2) = \frac{1}{n}\sum_{i=1}^{n}\exp\{-\beta g_i(\sigma^2)\}(g_i(\sigma^2) - 1) + \frac{\beta}{(\beta+1)^{\frac{3}{2}}}. \tag{59}$$

将被更新。第 0 次迭代仅用于获取 $\widetilde{\boldsymbol{\theta}} = (\widetilde{\boldsymbol{\beta}}, \widetilde{\sigma}^2)^T$ 或 $\widetilde{\boldsymbol{\theta}}_{\beta=0} = (\widetilde{\boldsymbol{\beta}}_{\beta=0}, \widetilde{\sigma}^2_{\beta=0})^T$。第 i 次迭代包括更新$\mathbb{X}_n$、$\boldsymbol{D}_\beta(\widetilde{\boldsymbol{\theta}}_\beta)$ 和 $\boldsymbol{Y}(\widetilde{\boldsymbol{\theta}}_\beta)$，求解方程组（基于$\widetilde{\boldsymbol{\beta}}_{\beta,(i-1)}$,$\widetilde{\sigma}^2_{\beta,(i-1)}$）以获得$\widetilde{\boldsymbol{\beta}}_{\beta,(i)}$，更新$\widetilde{\sigma}^2_{\beta,(i)}$，通过求解仅含一个唯一变量的非线性方程来更新，并检查$\mathbb{X}_n^T(\widetilde{\boldsymbol{\theta}}_\beta)\left(\boldsymbol{Y}(\widetilde{\boldsymbol{\theta}}_\beta) - \mathbb{X}_n\boldsymbol{D}_\beta(\widetilde{\boldsymbol{\theta}}_\beta)\widetilde{\boldsymbol{\beta}}\right)$ 的范数是否小于预先指定的容差水平。如果是，则迭代序列停止；否则，继续迭代序列。根据前文所述的指示适当调整 Cholesky 分解后，可使用 R 的基础函数 `solve()` 高效求解线性方程组：

```
library(MASS)
XtX <- crossprod(X.tilde)
if (is.singular.matrix(XtX,tol=1e-08)){
print("singular cross-product matrix")
break
}
```

```
library(matrixcalc)
L <- t(chol(XtX))
beta.hat <- solve(t(L), solve(L, t(X.tilde) %*% Y.tilde))
```

利用 R 软件包 nleqslv （Klein 和 Sporleder（2020），版本 0.8-2，于 2020-12-01 安装）可以高效求解非线性方程，因为它能够以最少的迭代次数得到解。需要提供（59）和

$$\frac{\partial}{\partial\sigma^2}f(\sigma^2)=\frac{\beta}{n\sigma^2}\sum_{i=1}^{n}\exp\{-\beta g_i(\sigma^2)\}g_i^2(\sigma^2)-\frac{\beta+1}{n\sigma^2}\sum_{i=1}^{n}\exp\{-\beta g_i(\sigma^2)\}g_i(\sigma^2),$$

其中

$$\begin{aligned}g_i(\sigma^2)&=\frac{r_i^2(\boldsymbol{\beta})}{\sigma^2},\\ \frac{\partial}{\partial\sigma^2}g_i(\sigma^2)&=\left(-\frac{1}{\sigma^2}\right)\frac{r_i^2(\boldsymbol{\beta})}{\sigma^2}=\left(-\frac{1}{\sigma^2}\right)g_i(\sigma^2),\end{aligned}$$

使用以下代码:

```
get.eq <- function(s2) {
g <- (res*res)/s2
eg2 <- exp(-beta*g/2.0)
v <- (eg2) * (g-1.0) #(D.half.beta %*% D.half.beta) %*% (g-1.0)
v <- mean(v)+beta/((beta+1.0)^1.5)
return(v)
}
get.grad.eq <- function(s2) {
g <- (res*res)/s2
eg2 <- exp(-beta*g/2.0)
dv <- beta * (eg2) * (g*g) #(D.half.beta %*% D.half.beta) %*% (g-1.0)
dv <- dv - (beta+1.0) * (eg2) * g
dv <- mean(dv)/s2
return(dv)
}
library(nleqslv)
initialG <- c(sigma2.hat)
sp <- list(trace=1,allowSingular=TRUE,btol = 1.e-5)
zero <- nleqslv(initialG, get.eq, jac=get.grad.eq, method="Newton",control=sp)
sigma2.hat <- zero$x
```

E.2.2 Breusch-Pagan 和 Koenker 的 β-分数检验统计量

以与经典方法类似的方式，我们得到

$$R_n(\widetilde{\boldsymbol{\theta}}_\beta) = \frac{ESS(\widetilde{\boldsymbol{\theta}}_\beta)}{\frac{2(2\beta^2+1)}{(2\beta+1)^{5/2}} - \frac{\beta^2}{(\beta+1)^3}},$$
$$Q_n(\widetilde{\boldsymbol{\theta}}_\beta) = \frac{ESS(\widetilde{\boldsymbol{\theta}}_\beta)}{TSS(\widetilde{\boldsymbol{\theta}}_\beta)/n} = nR^2_{OLS}(\widetilde{\boldsymbol{\theta}}_\beta).$$

因此，一旦我们计算出 $\widetilde{\boldsymbol{\theta}}_\beta$，接下来采用的方案就与经典检验中完全相同。利用普通最小二乘法的经典方差分析即可轻松计算，因此可以使用 `ln()` 函数及其相关函数（来自 `R`）。

附录参考文献

[1] Davidson, R. 和 MacKinnon, J. G.（2021）。《计量经济学中的估计与推断》。牛津大学出版社，纽约。

[2] Klein, S. 和 Sporleder, C.（2020）。nleqslv：求解非线性方程及方程组。R 包，版本 0.8-2。 https://CRAN.R-project.org/package=nleqslv