2010 年度 学部 エコノメトリックス
第 17 回講義メモ (訂正版)

2010 年 12 月 7 日

4.2 攪乱項の分散不均一性 (続)

不均一分散の検出法

攪乱項 ϵ_i の分散が不均一であるかどうかは、通常は未知である。したがって、標本データおよび線形回帰モデルの推定結果から不均一分散の存在を検出する必要がある。

検出方法は、次の 3 つに大別できる。

(a) **図による方法**　残差の 2 乗 e_i^2 と説明変数 X_i との散布図を描く。e_i^2 が X_i に関わらずランダムに散らばっていれば均一分散。そうでなければ、不均一分散の可能性が疑われる。

(b) **データの特性から推測する方法**　標本データが複数のサブ・グループ

に分割できる場合、サブ・グループ毎に分散が異なっている可能性があるので、各サブ・グループにおける分散が同一か否かの仮説検定を行えばよい。

(c) **検定統計量による方法**　Goldfeld=Quandt 検定、Breusch=Pagan=Godfrey 検定、などが用いられる。

Goldfeld=Quandt 検定

$\sigma_i^2 \propto X_i$ という関係があると考えられる場合の検定法。まず、X_i を昇順（または降順）に並べ替え、データを $\frac{1}{2}(n-c)$ 個、c 個、$\frac{1}{2}(n-c)$ 個のサブ・グループに 3 分割する。そして、第 1 のグループと第 3 のグループのそれぞれについて OLS 推定し、その攪乱項の推定量を $\hat\sigma_j^2$ $(j = 1, 3)$ で

表すと、

$$\begin{cases} H_0: & \sigma_1^2 = \sigma_3^2 \\ H_1: & \sigma_1^2 \neq \sigma_3^2 \end{cases}$$

という仮説において、帰無仮説 H_0 の下では

$$\frac{\hat\sigma_1^2}{\hat\sigma_3^2} \sim F\Big(\frac{n-c}{2} - k, \frac{n-c}{2} - k\Big)$$

となるので、これによって検定できる。

Breush=Pagan=Godfrey 検定

攪乱項 ϵ_i の分散が、外生変数 $\underset{m\times 1}{\boldsymbol{z}_i} = (z_{1,i}\ z_{2,i}\ \cdots\ z_{m,i})'$ の一次結合

$$\mathrm{Var}(\epsilon_i) = \sigma_i^2 = \alpha_o + \boldsymbol{z}_i'\boldsymbol{\alpha} = \alpha_0 + \sum_{j=1}^m \alpha_j z_{j,i}$$

ただし、$\boldsymbol{\alpha} = (\alpha_1\ \alpha_2\ \cdots\ \alpha_m)'$、として表現できるときの検定である。

均一分散であれば、$\boldsymbol{\alpha} = \boldsymbol{0}$ であるので、仮説を

$$\begin{cases} H_0 & : & \boldsymbol{\alpha} = \boldsymbol{0} \\ H_1 & : & \boldsymbol{\alpha} \neq \boldsymbol{0} \end{cases}$$

とおいて検定する。攪乱項 ϵ_i の分散のモデルを扱いやすいように次のように変形しよう。

$$\mathrm{Var}(\epsilon_i) = \sigma_i^2 = \sigma^2(1 + \boldsymbol{z}_i'\boldsymbol{\delta})$$

ただし、$\boldsymbol{\delta} = (\delta_1\ \delta_2\ \cdots\ \delta_m)'$ かつ $\sigma^2 = \alpha_0$, $\delta_i = \frac{\alpha_i}{\alpha_0}$ である。$\boldsymbol{\delta}$ は未知であるから、次の回帰モデルで推定する。

$$q_i \equiv \frac{e_i^2}{\sigma^2} - 1 = \boldsymbol{z}_i'\boldsymbol{\delta}$$

均一分散であれば、$\boldsymbol{\delta} = \boldsymbol{0}$ であるので、仮説を

$$\begin{cases} H_0 & : & \boldsymbol{\delta} = \boldsymbol{0} \\ H_1 & : & \boldsymbol{\delta} \neq \boldsymbol{0} \end{cases}$$

とおいて検定する。上記の回帰モデルで、$\boldsymbol{\delta}$ の最小 2 乗推定量 $\boldsymbol{d}$ の分散は、

$$\mathrm{Var}(\boldsymbol{d}) = 2(\boldsymbol{Z}'\boldsymbol{Z})^{-1}, \qquad \underset{n\times m}{\boldsymbol{Z}} = \begin{pmatrix} \boldsymbol{z}_1' \\ \vdots \\ \boldsymbol{z}_n' \end{pmatrix}$$

となるから[*1]、$\boldsymbol{q} = (q_1 \ \cdots \ q_n)'$ とおくと

$$\frac{1}{2}\boldsymbol{d}'(\boldsymbol{Z}'\boldsymbol{Z})\boldsymbol{d} = \frac{1}{2}\boldsymbol{q}'\boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{q}$$

は帰無仮説の下で、自由度 m の χ^2 分布にしたがうので、これを用いて検定を行う。

実際には、σ^2 は未知であるので、一致推定量 $\tilde{\sigma}^2 = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} e_i^2$ で置きかえる。

*1 ϵ_i が正規分布に従うとき、すなわち $\epsilon_i \sim N(0, \sigma^2)$ のとき、$\rho \equiv \frac{\epsilon_i^2}{\sigma^2} \sim \chi^2(1)$ になる。このとき、$\chi^2(1)$ の性質より $\mathrm{E}(\rho) = 1$ かつ $\mathrm{Var}(\rho) = 2$。OLS 残差 e_i は ϵ_i の一致推定量であるから、$n \to \infty$ のとき、$\mathrm{E}(q_i) = 0$ かつ $\mathrm{Var}(q_i) = 2$ になることを使っている。

モデルに依存しない OLS 推定量の分散の推定法

今まで見てきたように、GLS 推定が行えるのは、攪乱項の分散に関して何らかのモデル化が行えるケースに限られる。では、モデル化が出来ない場合にはどうすればよいであろうか。

攪乱項の分散が不均一の場合でも、OLS 推定量は不偏推定量である。したがって、OLS 推定量の分散を正しく推定できさえすれば、（検出力は弱いが）サイズが正しい統計的検定を行うことができる。

この時に用いられるのが、White の不均一分散・一致分散共分散行列推定量 (HCCME: heteroscedasticity-Consistent Covariance Matrices Estimator) である。

いま、線形回帰モデルを $y_i = \tilde{\boldsymbol{x}}_i'\boldsymbol{\beta} + \epsilon_i$, $(i = 1, \ldots, n)$ で表すことにしよう。このとき、

$$\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X} = (\tilde{\boldsymbol{x}}_1, \cdots, \tilde{\boldsymbol{x}}_n)\begin{pmatrix} \sigma_1^2 & & \boldsymbol{O} \\ & \ddots & \\ \boldsymbol{O} & & \sigma_n^2 \end{pmatrix}\begin{pmatrix} \tilde{\boldsymbol{x}}_1' \\ \vdots \\ \tilde{\boldsymbol{x}}_n' \end{pmatrix}$$

を次式で推定することを考える。

$$\sum_{i=1}^{n} \tilde{\boldsymbol{x}}_i \tilde{\boldsymbol{x}}_i' e_i^2 \tag{4.2.1}$$

したがって、$\mathrm{Var}(\boldsymbol{b}_{OLS}) = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Omega}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}$ は、

$$\mathrm{Var}(\widehat{\boldsymbol{b}_{OLS}}) = (\boldsymbol{X}'\boldsymbol{X})^{-1}\left(\sum_{i=1}^{n} \tilde{\boldsymbol{x}}_i \tilde{\boldsymbol{x}}_i' e_i^2\right)(\boldsymbol{X}'\boldsymbol{X})^{-1} \tag{4.2.2}$$

で推定できる。これが White の HCCME である。