統計数理研究所彙報
第29巻 第2号 (1982)

誤差分散の巾乗不均一性と二段階最小二乗推定

統計数理研究所 鈴 木 義 一 郎

(1982年1月 受付)

1. はじめに

回帰モデルにおける誤差項が分散均一性の仮定を満さない場合を扱った文献は，大きく分けて二つある．その一つは，いくつかの同一グループ内では分散が一定であるが異るグループ間では違っているような場合で，[1], [2], [3] といった論文で推定問題などが種々議論されている．もう一つは，分散の値が説明変数の大きさに関係して変動するような場合で，典型的な例として分散が説明変数の二乗に比例して変わるようなケースが扱われている（[4], [5] 参照)．

いずれにせよ，この種の不均一性が先験情報のようなものを用いて特定できる場合には，適当な変数変換を用いて分散の値を均一化してから通常の最小二乗法 (OLS) を用いて回帰係数を推定すればよい．しかし分散の不均一性が予想されても，その正確な構造を決定すべき情報が得られない場合には，まず第一段階でその構造の推定を行い，次に回帰係数を推定するといった，いわゆる二段階最小二乗推定法を適用することが考えられる．

誤差分散が均一であるか，説明変数の二乗に比例するかを識別する実際的な手続きは，通常の最小二乗法を適用して得られた残差系列をプロットしてみて，分析者の直感的な判断によるケースが殆んどである．しかしこの種の判断は，データの性格であるとか，分析者の経験の多少などにより，必ずしも妥当な結論が得られない場合もある．

そこでこの小稿では，“誤差分散が説明変数の $2g$ 乗に比例する，ただし g の値は0と1の間にある”という仮定群を正しいものとして，通常の最小二乗法による残差系列の情報から g の値を推定し，その推定値に基づいたウエイトによる加重最小二乗推定を行うという，二段階最小二乗推定法に関する統計的性質を調べてみた．

誤差分散の構造が

$$\sigma^2 \cdot \mathrm{diag}(x_1^{2g}, x_2^{2g}, \cdots, x_n^{2g})$$

という形で σ^2 の値が既知と考えると，通常の最小二乗法による残差系列の情報を用いての，ここで提示した g の推定法はかなり満足すべき結果の得られることがわかる．また σ^2 の値も未知とした場合でも，g の値を既知として適用できる最良の加重最小二乗推定と，それほど見劣りすることはないという事実が，計算機による数値実験の結果明らかになった．

ともあれここで提示した，誤差分散の不均一性の構造を探知しながら行う二段階最小二乗推定法は，所与のデータから自動的に推定量を与えることができ，しかもその安定性は分散不均一性の構造に関する知識を利用した最良の加重推定量と比較しても遜色がないということなので，その利用法が別して単純であることから，実際に推奨に価する方法であることを強調しておきたい．

2. 不均一な誤差分散を有する回帰モデルと近似加重最小二乗推定量

次のような単回帰モデルを考える.

(1) $$\boldsymbol{y}=X\boldsymbol{\beta}+\varepsilon$$

ここで

$$\boldsymbol{y}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix},\qquad X=\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix},\qquad \varepsilon=\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix},\qquad \boldsymbol{\beta}=\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

$$\mathrm{E}\{\varepsilon\}=\mathbf{0},$$

$$\mathrm{V}(\varepsilon)=\mathrm{E}\{\varepsilon\varepsilon'\}=\sigma^2\Omega_g,\qquad \Omega_g=\begin{bmatrix} {x_1}^{2g} & & & 0 \\ & {x_2}^{2g} & & \\ & & \ddots & \\ 0 & & & {x_n}^{2g} \end{bmatrix}.$$

よく知られているように，このモデルに対する β の加重最小二乗 (WLS) 推定量は

(2) $$\tilde{\boldsymbol{\beta}}_g=(X'{\Omega_g}^{-1}X)^{-1}X'{\Omega_g}^{-1}\boldsymbol{y}$$

によって与えられ，さらに分散共分散行列は

(3) $$V_g(\tilde{\boldsymbol{\beta}}_g)=\sigma^2(X'{\Omega_g}^{-1}X)^{-1}$$

で与えられる．しかしこのような推定量 $\tilde{\boldsymbol{\beta}}_g$ は g の値が既知でないと構成することはできない．

そこで，g の近似値 h が利用できるものとして

(4) $$\tilde{\boldsymbol{\beta}}_h=(X'{\Omega_h}^{-1}X)^{-1}X'{\Omega_h}^{-1}\boldsymbol{y}$$

を推定量として用いたらどうだろうか．このタイプの推定量を便宜上近似加重最小二乗 (AWLS) 推定量と呼ぶことにしよう．容易にわかるように，AWLS 推定量の分散共分散行列は

(5) $$V_g(\tilde{\boldsymbol{\beta}}_h)=\sigma^2(X'{\Omega_h}^{-1}X)^{-1}X'{\Omega_h}^{-1}\Omega_g{\Omega_h}^{-1}X(X'{\Omega_h}^{-1}X)^{-1}$$

で与えられる．$\tilde{\boldsymbol{\beta}}_h$ の $\tilde{\boldsymbol{\beta}}_g$ に対する相対効率を，一般化分散の逆数の比

$$E_g(\tilde{\boldsymbol{\beta}}_h)=\frac{|V_g(\tilde{\boldsymbol{\beta}}_g)|}{|V_g(\tilde{\boldsymbol{\beta}}_h)|}$$

によって測ることにすれば

(6) $$E_g(\tilde{\boldsymbol{\beta}}_h)=\frac{|X'{\Omega_h}^{-1}X|^2}{|X'{\Omega_h}^{-1}\Omega_g{\Omega_h}^{-1}X|\,|X'{\Omega_g}^{-1}X|}=\frac{D(-h)^2}{D(-g)D(g-2h)}$$

のようにあらわされる．ここで

$$D(\alpha)=S_{2\alpha}S_{2+2\alpha}-{S_{1+2\alpha}}^2,$$

$$S_\alpha=\sum_{i=1}^{n}{x_i}^{\alpha}.$$

また，推定量 $\tilde{\boldsymbol{\beta}}_h$ の通常の最小二乗 (OLS) 推定量に対する相対効率は

$${E_g}^0(\tilde{\boldsymbol{\beta}}_h)=\frac{|V_g(\tilde{\boldsymbol{\beta}}_0)|}{|V_g(\tilde{\boldsymbol{\beta}}_h)|}=\frac{D(g)D(-h)^2}{D(0)^2D(g-2h)}$$

さらに，g の値を1と考えた加重最小二乗推定量 $\tilde{\boldsymbol{\beta}}_1$ に対する相対効率も

表 1 $\tilde{\boldsymbol{\beta}}_h$ の WLS 推定量 ($\tilde{\boldsymbol{\beta}}_g$) に対する相対効率

g \ h	0.0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
0.0	1.00	0.99	0.97	0.93	0.88	0.83	0.76	0.70	0.63	0.56	0.50	0.44	0.39	0.34	0.29	0.25	0.21	0.17	0.14	0.12	0.09
0.05	0.99	1.00	0.99	0.97	0.93	0.88	0.83	0.76	0.70	0.63	0.57	0.51	0.45	0.39	0.34	0.30	0.25	0.21	0.18	0.15	0.12
0.10	0.97	0.99	1.00	0.99	0.97	0.93	0.89	0.83	0.77	0.71	0.64	0.58	0.52	0.46	0.40	0.35	0.30	0.26	0.22	0.18	0.15
0.15	0.93	0.97	0.99	1.00	0.99	0.97	0.93	0.89	0.83	0.77	0.71	0.65	0.59	0.53	0.47	0.41	0.36	0.31	0.26	0.22	0.19
0.20	0.89	0.93	0.97	0.99	1.00	0.99	0.97	0.93	0.89	0.84	0.78	0.72	0.66	0.59	0.53	0.47	0.42	0.37	0.32	0.27	0.23
0.25	0.83	0.89	0.93	0.97	0.99	1.00	0.99	0.97	0.94	0.89	0.84	0.78	0.72	0.66	0.60	0.54	0.48	0.43	0.37	0.32	0.27
0.30	0.77	0.83	0.88	0.93	0.97	0.99	1.00	0.99	0.97	0.94	0.89	0.84	0.79	0.73	0.67	0.61	0.55	0.49	0.43	0.38	0.33
0.35	0.70	0.76	0.83	0.88	0.93	0.97	0.99	1.00	0.99	0.97	0.94	0.90	0.85	0.79	0.74	0.68	0.62	0.56	0.50	0.44	0.38
0.40	0.63	0.69	0.76	0.82	0.88	0.93	0.97	0.99	1.00	0.99	0.97	0.94	0.90	0.85	0.80	0.74	0.68	0.62	0.56	0.50	0.44
0.45	0.55	0.62	0.69	0.76	0.82	0.88	0.93	0.97	0.99	1.00	0.99	0.97	0.94	0.90	0.86	0.80	0.75	0.69	0.63	0.57	0.51
0.50	0.49	0.55	0.62	0.69	0.76	0.83	0.89	0.93	0.97	0.99	1.00	0.99	0.97	0.94	0.91	0.86	0.81	0.75	0.69	0.63	0.57
0.55	0.42	0.48	0.55	0.62	0.69	0.76	0.83	0.89	0.94	0.97	0.99	1.00	0.99	0.97	0.95	0.91	0.86	0.81	0.76	0.70	0.64
0.60	0.36	0.41	0.48	0.54	0.62	0.69	0.76	0.83	0.89	0.94	0.97	0.99	1.00	0.99	0.98	0.95	0.91	0.87	0.82	0.76	0.70
0.65	0.30	0.35	0.41	0.47	0.54	0.62	0.69	0.76	0.83	0.89	0.94	0.97	0.99	1.00	0.99	0.98	0.95	0.91	0.87	0.82	0.76
0.70	0.25	0.30	0.35	0.41	0.47	0.54	0.62	0.70	0.77	0.84	0.89	0.94	0.97	0.99	1.00	0.99	0.98	0.95	0.91	0.87	0.82
0.75	0.21	0.25	0.30	0.35	0.41	0.48	0.55	0.62	0.70	0.77	0.84	0.90	0.94	0.97	0.99	1.00	0.99	0.98	0.95	0.92	0.87
0.80	0.17	0.21	0.25	0.29	0.35	0.41	0.48	0.55	0.63	0.71	0.78	0.84	0.90	0.94	0.98	0.99	1.00	0.99	0.98	0.95	0.92
0.85	0.14	0.17	0.20	0.24	0.29	0.35	0.41	0.48	0.56	0.63	0.71	0.78	0.85	0.90	0.95	0.98	0.99	1.00	0.99	0.98	0.95
0.90	0.11	0.14	0.17	0.20	0.24	0.29	0.35	0.42	0.49	0.56	0.64	0.72	0.79	0.85	0.91	0.95	0.98	0.99	1.00	0.99	0.98
0.95	0.09	0.11	0.13	0.16	0.20	0.24	0.30	0.35	0.42	0.49	0.57	0.65	0.72	0.79	0.86	0.91	0.95	0.98	0.99	1.00	0.99
1.00	0.07	0.09	0.11	0.13	0.16	0.20	0.25	0.30	0.36	0.43	0.50	0.58	0.66	0.73	0.80	0.86	0.91	0.95	0.98	0.99	1.00

表 2 $\tilde{\boldsymbol{\beta}}_h$ の OLS 推定量 ($\tilde{\boldsymbol{\beta}}_0$) に対する相対効率

g \ h	0.0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
0.0	1.00	0.99	0.97	0.93	0.88	0.83	0.76	0.70	0.63	0.56	0.50	0.44	0.39	0.34	0.29	0.25	0.21	0.17	0.14	0.12	0.09
0.05	1.00	1.01	1.00	0.98	0.94	0.89	0.83	0.77	0.71	0.64	0.57	0.51	0.45	0.40	0.35	0.30	0.25	0.21	0.18	0.15	0.12
0.10	1.00	1.02	1.03	1.02	1.00	0.96	0.91	0.85	0.79	0.73	0.66	0.60	0.53	0.47	0.42	0.36	0.31	0.27	0.22	0.19	0.16
0.15	1.00	1.04	1.06	1.07	1.06	1.04	1.00	0.95	0.89	0.83	0.76	0.69	0.63	0.56	0.50	0.44	0.38	0.33	0.28	0.24	0.20
0.20	1.00	1.05	1.09	1.12	1.13	1.12	1.09	1.05	1.00	0.94	0.88	0.81	0.74	0.67	0.60	0.54	0.47	0.41	0.36	0.30	0.26
0.25	1.00	1.07	1.12	1.17	1.20	1.20	1.20	1.17	1.13	1.07	1.01	0.94	0.87	0.80	0.73	0.65	0.58	0.51	0.45	0.39	0.33
0.30	1.00	1.08	1.15	1.22	1.27	1.30	1.31	1.30	1.27	1.22	1.17	1.10	1.03	0.95	0.88	0.80	0.72	0.64	0.56	0.49	0.43
0.35	1.00	1.09	1.18	1.27	1.34	1.39	1.42	1.43	1.42	1.39	1.35	1.29	1.22	1.14	1.06	0.97	0.88	0.80	0.71	0.63	0.55
0.40	1.00	1.11	1.21	1.32	1.41	1.49	1.55	1.59	1.60	1.59	1.55	1.50	1.44	1.36	1.28	1.19	1.09	1.00	0.90	0.80	0.71
0.45	1.00	1.12	1.24	1.37	1.49	1.59	1.68	1.75	1.79	1.80	1.79	1.75	1.70	1.63	1.54	1.45	1.35	1.24	1.13	1.02	0.91
0.50	1.00	1.13	1.27	1.42	1.56	1.70	1.82	1.92	2.00	2.04	2.06	2.04	2.00	1.94	1.86	1.77	1.66	1.55	1.43	1.30	1.18
0.55	1.00	1.14	1.30	1.47	1.64	1.81	1.97	2.11	2.23	2.31	2.36	2.38	2.36	2.32	2.25	2.16	2.05	1.93	1.80	1.66	1.52
0.60	1.00	1.15	1.33	1.51	1.71	1.92	2.12	2.31	2.47	2.61	2.70	2.76	2.78	2.76	2.71	2.63	2.53	2.41	2.27	2.11	1.95
0.65	1.00	1.16	1.35	1.56	1.79	2.03	2.27	2.51	2.74	2.93	3.09	3.20	3.26	3.29	3.27	3.21	3.12	3.00	2.85	2.69	2.51
0.70	1.00	1.17	1.37	1.60	1.86	2.14	2.43	2.73	3.02	3.28	3.51	3.69	3.82	3.90	3.92	3.90	3.83	3.73	3.59	3.42	3.22
0.75	1.00	1.18	1.40	1.65	1.93	2.25	2.59	2.95	3.31	3.66	3.97	4.25	4.46	4.61	4.70	4.73	4.70	4.63	4.50	4.33	4.13
0.80	1.00	1.19	1.42	1.69	2.00	2.36	2.75	3.18	3.62	4.06	4.48	4.86	5.19	5.44	5.62	5.73	5.76	5.73	5.64	5.48	5.28
0.85	1.00	1.20	1.44	1.73	2.07	2.47	2.92	3.41	3.95	4.49	5.04	5.55	6.01	6.40	6.70	6.92	7.04	7.08	7.04	6.93	6.75
0.90	1.00	1.21	1.46	1.77	2.14	2.58	3.08	3.65	4.28	4.95	5.63	6.30	6.93	7.49	7.96	8.32	8.58	8.74	8.79	8.74	8.60
0.95	1.00	1.22	1.48	1.81	2.21	2.68	3.24	3.89	4.62	5.42	6.27	7.12	7.96	8.73	9.41	9.98	10.43	10.74	10.93	10.99	10.93
1.00	1.00	1.22	1.50	1.85	2.27	2.78	3.40	4.14	4.98	5.92	6.94	8.01	9.09	10.13	11.08	11.92	12.62	13.17	13.56	13.79	13.87

表 3 $\tilde{\boldsymbol{\beta}}_h$ の $\tilde{\boldsymbol{\beta}}_1$ に対する相対効率

g \ h	0.0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
0.0	10.54	10.46	10.22	9.83	9.31	8.70	8.03	7.33	6.62	5.94	5.28	4.65	4.07	3.54	3.04	2.59	2.19	1.83	1.51	1.23	1.00
0.05	8.24	8.31	8.24	8.05	7.75	7.34	6.87	6.35	5.81	5.27	4.74	4.22	3.74	3.28	2.85	2.46	2.10	1.77	1.48	1.22	1.00
0.10	6.43	6.58	6.63	6.58	6.43	6.19	5.87	5.50	5.10	4.68	4.25	3.83	3.43	3.04	2.67	2.33	2.01	1.71	1.45	1.21	1.00
0.15	5.01	5.20	5.32	5.36	5.32	5.20	5.01	4.76	4.46	4.15	3.81	3.48	3.14	2.82	2.50	2.20	1.92	1.66	1.42	1.20	1.00
0.20	3.90	4.10	4.26	4.36	4.39	4.36	4.26	4.11	3.91	3.67	3.42	3.15	2.88	2.61	2.34	2.09	1.84	1.61	1.39	1.18	1.00
0.25	3.03	3.23	3.40	3.54	3.62	3.65	3.62	3.54	3.41	3.25	3.06	2.86	2.64	2.42	2.20	1.98	1.76	1.55	1.36	1.17	1.00
0.30	2.35	2.54	2.71	2.86	2.97	3.04	3.07	3.04	2.98	2.88	2.74	2.59	2.42	2.24	2.06	1.87	1.68	1.50	1.33	1.16	1.00
0.35	1.82	1.99	2.16	2.31	2.44	2.53	2.59	2.61	2.60	2.54	2.46	2.35	2.22	2.08	1.93	1.77	1.61	1.45	1.30	1.15	1.00
0.40	1.41	1.56	1.72	1.86	1.99	2.11	2.19	2.24	2.26	2.24	2.20	2.12	2.03	1.92	1.80	1.68	1.54	1.41	1.27	1.13	1.00
0.45	1.10	1.23	1.36	1.50	1.63	1.75	1.84	1.92	1.96	1.97	1.96	1.92	1.86	1.78	1.69	1.59	1.48	1.36	1.24	1.12	1.00
0.50	0.85	0.96	1.08	1.20	1.33	1.44	1.55	1.64	1.70	1.74	1.75	1.74	1.70	1.65	1.58	1.50	1.41	1.32	1.21	1.11	1.00
0.55	0.66	0.75	0.86	0.97	1.08	1.19	1.30	1.39	1.47	1.52	1.56	1.57	1.56	1.53	1.48	1.43	1.35	1.28	1.19	1.10	1.00
0.60	0.51	0.59	0.68	0.78	0.88	0.98	1.09	1.18	1.27	1.34	1.39	1.42	1.43	1.42	1.39	1.35	1.30	1.23	1.16	1.08	1.00
0.65	0.40	0.46	0.54	0.62	0.71	0.81	0.91	1.00	1.09	1.17	1.23	1.28	1.30	1.31	1.30	1.28	1.24	1.20	1.14	1.07	1.00
0.70	0.31	0.36	0.43	0.50	0.58	0.66	0.76	0.85	0.94	1.02	1.09	1.15	1.19	1.21	1.22	1.21	1.19	1.16	1.11	1.06	1.00
0.75	0.24	0.29	0.34	0.40	0.47	0.54	0.63	0.71	0.80	0.89	0.96	1.03	1.08	1.12	1.14	1.15	1.14	1.12	1.09	1.05	1.00
0.80	0.19	0.23	0.27	0.32	0.38	0.45	0.52	0.60	0.69	0.77	0.85	0.92	0.98	1.03	1.06	1.08	1.09	1.09	1.07	1.04	1.00
0.85	0.15	0.18	0.21	0.26	0.31	0.37	0.43	0.51	0.58	0.67	0.75	0.82	0.89	0.95	0.99	1.03	1.04	1.05	1.04	1.03	1.00
0.90	0.12	0.14	0.17	0.21	0.25	0.30	0.36	0.42	0.50	0.58	0.66	0.73	0.81	0.87	0.93	0.97	1.00	1.02	1.02	1.02	1.00
0.95	0.09	0.11	0.14	0.17	0.20	0.25	0.30	0.36	0.42	0.50	0.57	0.65	0.73	0.80	0.86	0.91	0.95	0.98	1.00	1.01	1.00
1.00	0.07	0.09	0.11	0.13	0.16	0.20	0.25	0.30	0.36	0.43	0.50	0.58	0.66	0.73	0.80	0.86	0.91	0.95	0.98	0.99	1.00

$$E_g{}^1(\tilde{\boldsymbol{\beta}}_h)=\frac{|V_g(\tilde{\boldsymbol{\beta}}_1)|}{|V_g(\tilde{\boldsymbol{\beta}}_h)|}=\frac{D(g-2)D(-h)^2}{D(-1)^2D(g-2h)}$$

のように与えられる．

表 1，表 2 および表 3 に，$n=20$，$x_i=i$ の場合について，$\tilde{\boldsymbol{\beta}}_h$ のこの種の相対効率を示してみた．

$$E_g(\tilde{\boldsymbol{\beta}}_h) \geqq 0.9, \qquad E_g{}^0(\tilde{\boldsymbol{\beta}}_h) \geqq 1, \qquad E_g{}^1(\tilde{\boldsymbol{\beta}}_h) \geqq 1$$

といった不等式を満足する h の範囲が存外広いし，他の n についても同じような傾向を示すことから，AWLS 推定量の利用可能性が十分示唆される．

3. 近似的加重最小二乗法による残差の統計的性質

回帰モデル(1)に対して(4)の AWLS 推定量を用いたときの残差平方和は

$$\boldsymbol{e}_h'\boldsymbol{e}_h = (\boldsymbol{y}-X\tilde{\boldsymbol{\beta}}_h)'(\boldsymbol{y}-X\tilde{\boldsymbol{\beta}}_h) = \boldsymbol{y}P_h'P_h\boldsymbol{y}$$

で与えられる．ここで

$$P_h = I - X(X'\Omega_h{}^{-1}X)^{-1}X'\Omega_h{}^{-1}$$

は $h=0$ の場合を除けばベキ等行列ではない．

まず，行列 $X(X'\Omega_h{}^{-1}X)^{-1}X'$ の (i, j) 要素が

(7) $$z_{ij}{}^{(h)} = \frac{S_{2-2h}-(x_i+x_j)S_{1-2h}+x_ix_jS_{-2h}}{S_{-2h}S_{2-2h}-S_{1-2h}{}^2}$$

と表わされるから，P_h の (i, j) 要素も

$$\delta_{ij}-z_{ij}{}^{(h)}x_j{}^{-2h}$$

のように表わされる．したがって

$$\begin{aligned}
\mathrm{tr}\,[P_h'P_h\Omega_g] &= \sum_{j=1}^{n} x_j{}^{2g}\sum_{i=1}^{n}(\delta_{ij}-z_{ij}{}^{(h)}x_j{}^{-2h})^2 \\
&= \sum_{j=1}^{n}(1-2z_{jj}{}^{(h)}x_j{}^{-2h})x_j{}^{2g}+\sum_{j=1}^{n}x_j{}^{2g}\sum_{i=1}^{n}z_{ij}{}^{(h)2}x_j{}^{-4h} \\
&= S_{2g}-2\frac{A_h}{K_h}+\frac{B_hS_0-2C_hS_1+D_hS_2}{K_h{}^2} \\
&\equiv m_h(g)
\end{aligned}$$

ここで

$$\begin{aligned}
A_h &= S_{2-2h}S_{2g-2h}-2S_{2-2h}S_{1-2h}S_{1+2g-4h}+S_{1-2h}{}^2S_{2+2g-4h} \\
B_h &= S_{2-2h}{}^2S_{2g-4h}-2S_{2-2h}S_{1-2h}S_{1+2g-4h}+S_{1-2h}{}^2S_{2+2g-4h} \\
C_h &= S_{2-2h}S_{1-2h}S_{2g-4h}-(S_{2-2h}S_{-2h}+S_{1-2h}{}^2)S_{1+2g-4h}+S_{1-2h}S_{-2h}S_{2+2g-4h} \\
D_h &= S_{1-2h}{}^2S_{2g-4h}-2S_{1-2h}S_{-2h}S_{1+2g-4h}+S_{-2h}{}^2S_{2+2g-4h} \\
K_h &= S_{-2h}S_{2-2h}-S_{1-2h}{}^2.
\end{aligned}$$

かくて残差平方和の期待値は

$$\begin{aligned}
\mathrm{E}\{\boldsymbol{e}_h'\boldsymbol{e}_h\} &= \mathrm{E}\{\boldsymbol{y}'P_h'P_h\boldsymbol{y}\} \\
&= \mathrm{tr}[P_h'P_h\mathrm{E}\{\boldsymbol{y}\boldsymbol{y}'\}] \\
&= \mathrm{tr}[P_h'P_h\mathrm{E}\{\varepsilon\varepsilon'\}] \\
&= \sigma^2\mathrm{tr}[P_h'P_h\Omega_g] \\
&= \sigma^2 m_h(g)
\end{aligned}$$

のように算出される．特に $h=0$ の場合には

$$B_0S_0-2C_0S_1+D_0S_2 = K_0A_0$$

であるから，

$$m_0(g)=S_{2g}-\frac{A_0}{K_0}=S_{2g}-\frac{S_2S_{2g}-2S_1S_{1+2g}+S_1^2S_{2+2g}}{S_0S_2-S_1{}^2}\ .$$

さらに，誤差ベクトル ε が正規分布 $N(0,\sigma^2\Omega_g)$ に従う場合には，推定値 $\tilde{\boldsymbol{\beta}}_h$ による残差ベクトル $\boldsymbol{e}_h$ も正規分布 $N(0,\sigma^2P_h\Omega_gP_h')$ に従うから

$$\mathrm{E}\{e_{hi}{}^2\}=\sigma^2v_{hi},\qquad \mathrm{E}\{|e_{hi}|\}=\sigma\sqrt{\frac{2}{\pi}v_{hi}}$$

という関係が成り立つ．ここで v_{hi} は行列 $P_h\Omega_gP_h$ の (i,i) 要素で

$$v_{hi}=(1-2z_{00}{}^{(h)}x_i{}^{-2h})x_i{}^{2g}+\sum_{j=1}^{n}z_{ij}{}^{(h)2}x_j{}^{2g-4h}$$

で与えられる ($z_{ij}{}^{(h)}$ は (7) で与えられる)．

特に $h=0$ の場合には

$$v_{0i}=M_{n,i}(g) \tag{8}$$

で与えられる．ここで

$$M_{n,i}(\alpha)=\left(1-2\frac{I_i-x_iJ_i}{K_0}\right)x_i{}^{2\alpha}+\frac{I_i{}^2S_{2\alpha}-2I_iJ_iS_{1+2\alpha}+J_i{}^2S_{2+2\alpha}}{K_0{}^2}$$
$$I_i=S_2'-x_iS_1'$$
$$J_i=S_1-x_iS_0'$$
$$K_0=S_0S_2'-S_1'^2.$$

4. 巾乗の推定と推定二段階最小二乗 (ETSLS) 推定量

まず σ^2 の値が既知である場合について，巾乗パラメータ g の推定問題を考える．$\sigma=1$ と仮定して一般性を失わない．最初に，通常の最小二乗推定量

$$\tilde{\boldsymbol{\beta}}_0=(X'X)^{-1}X'\boldsymbol{y}$$

を用いて残差平方和

$$\boldsymbol{e}_0'\boldsymbol{e}_0=(\boldsymbol{y}-X\tilde{\boldsymbol{\beta}}_0)'(\boldsymbol{y}-X\tilde{\boldsymbol{\beta}}_0)$$

を算出する．第2節でみたように，$h=0$ の場合の残差ベクトルは正規分布 $N(\mathbf{0},P_0\Omega_gP_0')$ に従う．ここで $P_0\Omega_gP_0'$ の第 (i,i) 要素は (8) で与えられる v_{0i} になる．これより，ベクトル $\boldsymbol{e}_0$ の第 i 要素 e_{0i} に対して

$$\mathrm{E}\{e_{0i}{}^2\}=v_{0i},\qquad \mathrm{E}\{|e_{0i}|\}=\sqrt{\frac{2}{\pi}v_{0i}}$$

といった関係が成立する．

そこで，$h\geqq0$ に対して

$$M_1(h)=\sum_{i=1}^{n}\left(|e_{0i}|-\sqrt{\frac{2}{\pi}M_{n,i}(h)}\right)^2$$

と置いて

$$\min_h M_1(h)=M_1(\hat{g}_1)$$

となる $\hat{g}_1$ が g に対する一つの推定量を与えることがわかる．

また回帰係数ベクトル $\boldsymbol{\beta}$ に対する推定量としては，g の代わりにこの $\hat{g}_1$ を用いた二段階最

小二乗 (TSLS) 推定量

$$\hat{\boldsymbol{\beta}}_1=\tilde{\boldsymbol{\beta}}_{g_1}=(X'\Omega_{\hat{g}_1}{}^{-1}X)^{-1}X'\Omega_{\hat{g}_1}{}^{-1}\boldsymbol{y}$$

を利用することができる．ただこの推定量は，$\hat{g}_1$ と $\boldsymbol{y}$ の独立性が保証されないから，$\boldsymbol{\beta}$ に対する不偏推定量ではない．しかし後の数値実験の結果をみると，その偏りはほとんど気にするほどのものではないことが実証できる．

表 4 $\hat{g}_1$ の度数分布

g \ $\hat{g}_1$	0.0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
0.0	631	236	102	25	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.05	369	259	239	100	27	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.10	177	196	255	238	97	31	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.15	69	104	204	245	242	97	33	6	0	0	0	0	0	0	0	0	0	0	0	0	0
0.20	25	42	109	206	237	245	96	34	6	0	0	0	0	0	0	0	0	0	0	0	0
0.25	10	16	38	116	206	234	240	101	34	5	0	0	0	0	0	0	0	0	0	0	0
0.30	2	9	16	37	123	199	227	245	102	34	6	0	0	0	0	0	0	0	0	0	0
0.35	0	2	9	17	37	127	193	230	244	100	35	6	0	0	0	0	0	0	0	0	0
0.40	0	0	2	8	20	37	125	196	226	240	105	35	6	0	0	0	0	0	0	0	0
0.45	0	0	0	2	8	20	39	126	192	225	238	109	34	7	0	0	0	0	0	0	0
0.50	0	0	0	0	2	8	22	38	128	189	225	239	109	33	7	0	0	0	0	0	0
0.55	0	0	0	0	0	2	9	22	39	130	193	217	238	110	33	7	0	0	0	0	0
0.60	0	0	0	0	0	0	1	10	23	42	125	195	217	233	112	35	7	0	0	0	0
0.65	0	0	0	0	0	0	0	1	10	22	43	126	196	213	230	113	39	6	1	0	0
0.70	0	0	0	0	0	0	0	0	2	9	23	45	123	203	211	221	113	42	7	1	0
0.75	0	0	0	0	0	0	0	0	0	3	8	23	45	125	200	214	214	118	41	8	1
0.80	0	0	0	0	0	0	0	0	0	0	3	8	24	51	124	196	213	213	117	43	8
0.85	0	0	0	0	0	0	0	0	0	0	0	3	7	26	53	123	191	217	209	120	51
0.90	0	0	0	0	0	0	0	0	0	0	0	0	3	7	26	54	126	191	211	211	171
0.95	0	0	0	0	0	0	0	0	0	0	0	0	0	3	8	26	54	130	189	208	382
1.00	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	8	30	52	129	192	586

表4に，$n=20$, $x_i=i$ の場合の 1000 回の数値実験による $\hat{g}_1$ の度数分布を示してある．ただし推定量を構成する場合には，g の変化する範囲が 0 と 1 の間に制限してあるので，当然 $\hat{g}_1$ は g の不偏推定ではない．

次に σ^2 の値が未知の場合を考えよう．通常の最小二乗推定を用いたときの残差ベクトル $\boldsymbol{e}_0$ は，正規分布 $N(\boldsymbol{0},\sigma^2P_0\Omega_gP_0')$ に従うから，i 番目の要素に対して

$$\mathrm{E}\{e_{0i}{}^2\}=\sigma^2v_{0i},\qquad \mathrm{E}\{|e_{0i}|\}=\sigma\sqrt{\frac{2}{\pi}V_{0i}}$$

となる．この場合には

$$M_\sigma(h)=\sum_{i=1}^{n}\left(|e_{0i}|-\sigma\sqrt{\frac{2}{\pi}M_{n,i}(h)}\right)^2$$

と置くと，これは

$$\sigma(h)=\sum_{i=1}^{n}|e_{0i}|\sqrt{M_{n,i}(h)}\Bigg/\sqrt{\frac{2}{\pi}\sum_{i=1}^{n}M_{n\,i}(h)}$$

のときに最小値

$$M(h)=\sum_{i=1}^{n} e_{0i}^{2}-\left(\sum_{i=1}^{n}|e_{0i}|\sqrt{M_{n,i}(h)}\right)^{2}\Big/\sum_{i=1}^{n} M_{n,i}(h)$$

をとることがわかるから

$$\min_{h} M(h)=M(\hat{g})$$

なる $\hat{g}$ が巾乗 g に対する推定量を与える．従って $\boldsymbol{\beta}$ に対しても，この $\hat{g}$ を用いた推定二段階最小二乗推定量

$$\hat{\boldsymbol{\beta}}=\tilde{\boldsymbol{\beta}}_{\hat{g}}=(X'\Omega_{\hat{g}}^{-1}X)^{-1}X'\Omega_{\hat{g}}^{-1}\boldsymbol{y}$$

を用いることができる．表5に，$n=20$, $x_i=i$ の場合の1000回の数値実験による $\hat{g}$ の度数分布を示した．$\hat{g}_1$ の場合（σ 既知）と比較してかなりばらつきが大きくなる．

表　5

g ＼ $\hat{g}$	0.0	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
0.0	523	75	84	64	57	47	36	22	19	21	13	8	4	8	4	3	1	4	0	0	7
0.05	447	74	84	70	73	54	45	30	30	17	16	14	10	7	7	5	3	2	3	2	7
0.10	374	79	76	74	64	71	65	36	29	24	30	9	14	12	8	8	8	4	1	3	11
0.15	296	82	84	74	75	61	56	59	43	22	35	17	19	11	14	11	7	10	4	1	19
0.20	233	67	82	79	73	73	67	56	43	43	32	30	18	19	8	16	12	6	9	7	27
0.25	173	71	58	91	73	75	65	71	47	38	40	37	27	18	19	12	13	9	15	5	43
0.30	131	55	63	69	76	75	80	54	69	53	34	32	33	28	22	17	18	13	12	9	57
0.35	99	46	55	54	76	69	78	77	48	63	46	39	41	18	26	28	13	18	17	13	76
0.40	67	41	45	51	60	74	69	74	79	49	52	42	45	37	15	24	24	15	20	15	102
0.45	35	44	40	43	52	62	68	75	76	58	54	50	41	45	38	14	25	23	18	10	129
0.50	24	27	43	36	43	57	60	60	80	70	57	45	43	51	38	37	21	17	24	19	148
0.55	16	17	29	42	38	37	55	67	63	83	49	66	36	42	51	37	31	27	17	21	176
0.60	11	12	19	35	41	26	49	52	69	64	69	58	53	39	39	44	48	18	32	20	202
0.65	7	11	14	19	38	35	31	53	54	59	69	63	52	41	52	41	34	43	23	26	235
0.70	5	8	10	20	22	36	32	38	48	53	67	57	65	44	42	50	36	41	31	22	273
0.75	3	6	7	13	19	33	26	31	45	51	59	61	53	54	44	39	47	38	42	27	302
0.80	2	5	6	11	6	30	29	25	43	45	50	55	53	59	49	39	40	38	39	38	338
0.85	2	1	6	12	8	14	29	30	21	41	53	48	55	48	58	43	45	36	34	44	372
0.90	2	0	6	7	10	7	22	25	26	34	41	53	45	56	44	51	43	41	42	33	412
0.95	2	0	3	5	13	7	12	19	29	25	36	42	56	38	53	49	48	41	34	38	450
1.00	2	0	2	4	6	13	7	15	20	32	26	38	45	53	40	45	48	45	40	33	486

さらに σ^2 に対する推定量も

$$\hat{\sigma}^2=\boldsymbol{e}_0'\boldsymbol{e}_0/m_0(\hat{g})$$

で与えられるであろうことが予想される．しかしこの推定量は，通常の最小二乗推定による残差を用いたものであるから，二段階最小二乗推定量 $\hat{\beta}$ にもとずく残差

$$\boldsymbol{e}_{\hat{g}}'\boldsymbol{e}_{\hat{g}}=(\boldsymbol{y}-X\hat{\boldsymbol{\beta}})'(\boldsymbol{y}-X\hat{\boldsymbol{\beta}})$$

を考えたほうがよい．この期待値が，形式的には

$$\sigma^2\mathrm{E}\{\mathrm{tr}[P_{\hat{g}}'P_{\hat{g}}\Omega_{\hat{g}}]\}$$

といった値に近いことを予想して

$$\tilde{\sigma}^2 = \boldsymbol{e}_{\hat{g}}'\boldsymbol{e}_{\hat{g}}/m_{\hat{g}}(\hat{g})$$

という推定量を考えることができる．ここで

$$m_g(g) = S_{2g} - \frac{S_2S_{-2g} - 2S_1S_{1-2g} + S_0S_{2-2g}}{S_{-2g}S_{2-2g} - S_{1-2g}{}^2}.$$

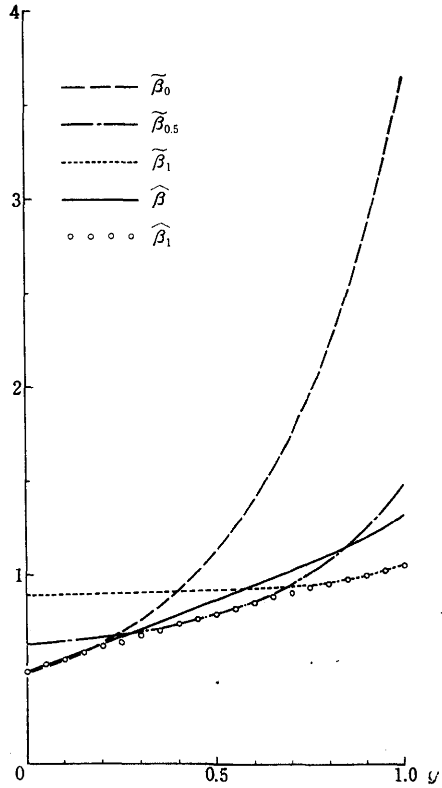

図 1 各種推定量の標準偏差
(切片（第1成分）の場合)

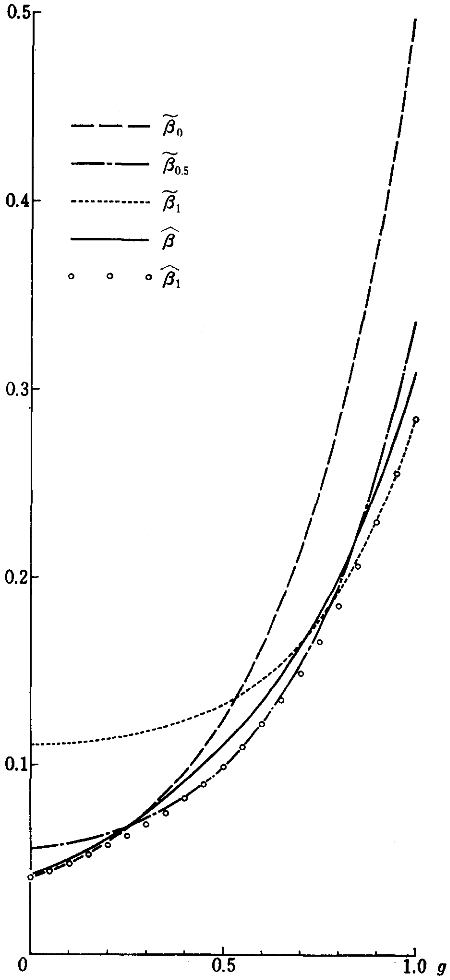

図 2 各種推定量の標準偏差
(勾配（第2成分）の場合)

図1と図2には，$n=20$, $x_i=i$ の場合の 1000 回の数値実験による5種類の推定量 $\hat{\beta}$, $\hat{\beta}_1$, $\tilde{\beta}_0$, $\tilde{\beta}_{0.5}$, $\tilde{\beta}_1$ の標準偏差の値を比較してみた．$\hat{\beta}_1$ の場合と同様，$\hat{g}$ が g の不偏推定でもなければ $\boldsymbol{y}$ と独立でもないから，β に対する不偏推定ではない．しかしその偏りはほとんど気にしなくてもよい程度のものである．総じて $\hat{\beta}_1$ がいいが，これは σ^2 の値が既知というやゝ非現実的な場合の推定量である．σ^2 の値が未知の場合の推定量 $\hat{\beta}$ は，g の値を 0.5 と仮定した $\tilde{\beta}_{0.5}$ とい

う推定量とある程度まで似通っている．当然予想されることではあるが，g の値の両端で $\hat{\beta}$ のほうがよく，真中の辺では $\tilde{\beta}_{0.5}$ がいい．また $\tilde{\beta}_0$ と $\tilde{\beta}_1$ を比較した場合，g の値が1に近づくにつれ前者が非常に不安定になるが，$\hat{\beta}_1$ のほうは全般に安定している．つまり，分散が説明変数の大きさに（必ずしも二乗とは限らず）比例して変わるような場合には，通常の最小二乗推定量 $\tilde{\beta}_0$ よりは，$\tilde{\beta}_{0.5}$ や $\tilde{\beta}_1$ を用いたほうがよい．しかしそれよりも自動的に利用できるという観点からも，ここで提示した推定二段階最小二乗推定量 $\hat{\beta}$ を使うほうが妥当である．標本数 n が他の値をとる場合についても同じような数値実験を試みた結果から，この結論は他の場合についても同様に成り立っていることがわかる．

参 考 文 献

[1] Hartley, H.O. and Jayatillake, K.S.E. (1973). Estimation for linear models with unequal variances, *J. Amer. Statist. Ass.*, **68**, 189–191.

[2] Rao, C.R. (1970). Estimation of heteroscedastic variances in linear models, *J. Amer. Statist. Ass.*, **65**, 161–172.

[3] Rao, C.R. and Subrahmaniam, K. (1971). Combining independent estimators and estimation in linear regression with unequal variances, *Biometrics*, **27**, 971–990.

[4] 佐和隆光 (1979). 回帰分析，朝倉書店.

[5] Seber, G.A.F. (1977). *Linear Regression Analysis*, John Wiley and Sons.

Power-Heteroscedasticity of Residual Variances and Two-Step Least Square

Giitiro Suzuki
(The Institute of Statistical Mathematics)

In this note, some modified Two-Step Least Square method are proposed for estimation of regression parameters in simple linear regression model with Power-Heteroscedastic variances. Some computer simulation shows that this method is very simple for use and nearly unbiased and so stable as the best linear unbiased estimator composed by using the parameter of power-heteroscedasticity.

The following simple regression model is considered:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

with

$$\mathrm{V}(\boldsymbol{\varepsilon}) = \sigma^2 \Omega_g, \qquad \Omega_g = \mathrm{diag}(x_1^{2g}, x_2^{2g}, \cdots, x_n^{2g})$$

where $g \in [0, 1]$ is unknown and σ^2 is known or unknown. Writing

$$\tilde{\boldsymbol{\beta}}_h = (X'\Omega_h^{-1}X)^{-1}X'\Omega_h^{-1}\boldsymbol{y},$$

then $\tilde{\boldsymbol{\beta}}_h$ is the best linear unbiased estimator of $\boldsymbol{\beta}$. Knowing the value of σ^2 (and assuming $\sigma^2=1$) we can get an estimator $\hat{g}_1$ minimizing

$$M_1(h) = \sum_{i=1}^{n} \left(|e_i| - \sqrt{\frac{2}{\pi} M_{n,i}(h)} \right)^2,$$

where e_i is the i-th component of residual vector obtained by $\tilde{\boldsymbol{\beta}}_0$ (ordinary least square estimator) and

$$M_{n,i}(h) = (1-2(I_i - x_i J_i)/K) x_i^{2h} + (I_i^2 S_{2h} - 2I_i J_i S_{1+2h} + J_i^2 S_{2+2h})/K^2.$$

$$\left(S_\alpha = \sum_{i=1}^{n} x_i^\alpha, \quad I_i = S_2 - x_i S_1, \quad J_i = S_1 - x_i S_0, \quad K = S_0 S_2 - S_1^2 \right)$$

Using the estimator $\hat{g}_1$ we can construct the Two-Step Least Square estimator

$$\hat{\boldsymbol{\beta}}_1 = \tilde{\boldsymbol{\beta}}_{\hat{g}_1}.$$

Similarily, when we don't know the value of σ^2, we can get an estimator $\hat{g}$ minimizing

$$M(h) = \sum_{i=1}^{n} e_i^2 - \frac{\sum_{i=1}^{n} |e_i| \sqrt{M_{n,i}(h)}}{\sum_{i=1}^{n} M_{n,i}(h)}$$

and then

$$\hat{\boldsymbol{\beta}} = \tilde{\boldsymbol{\beta}}_{\hat{g}}.$$