

# 構造を仮定しない不均一分散の推定

千木良 弘 朗\*  
斯波 恒 正†

## 要旨

本稿では、誤差項に不均一分散のある線形回帰モデルを考え、その不均一な分散の推定法を提案する。これまでは、不均一分散の推定は分散の構造をモデル化してから行われていた。これに対し本稿では、不均一分散の構造に一切の制約を置くことなくこれを推定する。これが、我々の推定法の特徴である。推定の基本的なアイデアは、White (1980) が提案した回帰パラメーター推定量の分散に対する一致推定量から、不均一分散に関する情報を抽出するというものである。White (1980) の推定量は不均一分散のモデル化を行わずに導出できるので、それを利用したこの推定量も制約を置かずに導出できる。モンテカルロ実験を行ったところ、推定量は適切に機能することが確かめられた。

## 1 はじめに

計量経済モデルにおいて、その誤差項の分散が観測値毎に一定ではないことがよくある。例えば、企業の利潤を考えると、利潤の変動は大規模な企業では大きく小規模な企業では小さいと思われる。よって、企業の利潤を説明する計量経済モデルの誤差項は企業規模に応じて分散が変わると考えられるのである。この現象は不均一分散と呼ばれ、不均一分散下での回帰係数の推定や検定、不均一分散性の検定等、多くの研究が行われてきた。それらの研究の中で、特に不均一分散下でのパラメーターの推定 (Generalized Least Squares (GLS) 推定) には不均一分散の推定が重要な役割を果たすことが知られている。また、多くの金融データにおいても、分散は時系列方向にも横断面方向にも不均一であると考えられて

\*日本学術振興会特別研究員 (一橋大学大学院経済学研究科)

†一橋大学大学院経済学研究科

いる。例えば、大規模な企業の株価と小規模な企業の株価の変動は異なるだろうし、同じ1つの企業でも今日の株価の変動と1ヵ月後の株価の変動は異なるだろう。ファイナンス、特に資産価格決定理論の分野では分散そのものが重要な役割を担うことが知られているので、この分野では回帰係数と共に不均一な分散そのものの推定に大きな関心が向けられている。

これまでの研究では、不均一分散の推定は、分散の構造をモデル化してから行われていた。先ほどの企業利潤の例では、よく誤差項の分散を企業規模の関数としてモデル化する。ファイナンスの分野でも、不均一分散の構造を表すのに、Auto Regressive Conditional Heteroscedasticity (ARCH) モデルや Stochastic Volatility (SV) モデル等、多くのモデルが提案されている。

これに対し本稿では、不均一分散の構造をモデル化することなく、それを推定する方法を提案する。このようにして、不均一分散の構造に一切の制約を置かないことが我々の推定法の特徴である。具体的には、誤差項の分散が不均一である線形回帰モデルを考え、その分散に制約を置くことなくそれを推定する。以下、推定の基本的なアイデアを述べる。White (1980) は不均一分散下で回帰係数の Ordinary Least Squares (OLS) 推定量の分散を一致推定しようと試み、Heteroscedasticity Consistent Estimator (HCE) を提案した。ここでHCEは不均一分散の構造をモデル化せず導かれていることが重要である。我々は、このHCEから誤差項の不均一な分散をパラメーターとする線形回帰モデルを導き、その回帰モデルのOLS推定量を以って不均一分散の推定量とする。HCEの導出には不均一分散のモデル化は必要ないので、それを利用した不均一分散の推定量も制約を置かずに導出できるのである。ただ、不均一分散の構造を仮定しないことから、推定すべきパラメーターの数が標本数と同じになることが容易に分かる。これより、不均一分散を推定する際には自由度が不足するであろうと推察できる。実際、この自由度不足の問題は深刻である。そこで、本稿ではある人工的な説明変数を使ってこの問題を回避する。モンテカルロ実験を行ったところ、この推定法は若干の問題もあるものの、適切に機能することが確かめられた。

我々の推定法は、新たな不均一分散性の検定や、不均一分散の構造を仮定しない頑健なGLS推定に応用できるだろう。また、モデルを仮定しない不均一分散

の推定法は、上に述べたように、ファイナンスの分野への貢献も期待される。

本稿の構成は以下の通りである。2節で、モデルと不均一分散の推定法を詳述する。3節では、2節で提案した推定法は自由度が不足して実行不能になる場合が多々あることを明らかにし、その問題への対処法を提案する。4節では、2、3節で提案した方法が適切に機能しているかをモンテカルロ実験で確かめる。最後に、5節でまとめと今後の展望を述べる。

## 2 モデルと推定法

標本数が $n$ 、説明変数の数が $K$ であり、誤差項に不均一分散がある以下の線形回帰モデルを考える。

$$y = X\beta + \varepsilon \quad (1)$$

ここで、 $y = (y_1, y_2, \dots, y_n)'$  は  $n \times 1$  の被説明変数ベクトル、 $X = (x_1, x_2, \dots, x_n)'$  は  $n \times K$  の外生的な説明変数行列、 $\beta$  は  $K \times 1$  のパラメーターベクトル、 $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)'$  は期待値0で分散 $\sigma^2 \Omega$ の誤差ベクトル、 $\sigma^2$ はスカラーのパラメーター、 $\Omega$ は $(w_1, w_2, \dots, w_n)$ を対角要素に持つ $n \times n$ の対角なパラメーター行列で、 $tr(\Omega) = n$ なる基準化を行う。

我々の目的は $\sigma^2 \Omega$ の推定である。これらのパラメーターの内、 $\sigma^2$ については、(1)をOLS推定して得られる $n \times 1$ の推定残差ベクトル $e = (e_1, e_2, \dots, e_n)'$ より一致推定量

$$\hat{\sigma}^2 = \frac{e'e}{n-K} \quad (2)$$

が得られる (Greene (2003) 等参照のこと)。従って、残るは $\Omega$ 、つまり

$$\omega = (w_1, w_2, \dots, w_n)' \subset \mathcal{R}_n^+$$

なる $n$ 個のパラメーターである。以下では $\omega$ の推定を考える。

モデル(1)のような不均一分散の下では、OLS推定量 $\hat{\beta}$ の極限分布の分散 $Var(\hat{\beta})$ は均一分散下でのそれとは異なることが知られている。White (1980) は $Var(\hat{\beta})$ を推定しようと試み、その一致推定量、いわゆるHCEを提案した。

$$\widehat{\text{Var}}(\hat{\beta}) = \left(\frac{X'X}{n}\right)^{-1} \hat{H} \left(\frac{X'X}{n}\right) \quad (3)$$

ここで、 $\hat{H} = \frac{1}{n} \sum_{i=1}^n x_i x_i' e_i^2 / n$ である。White (1980) は、適切な仮定の下で

$$\hat{H} \xrightarrow{p} H = \lim \frac{1}{n} \sum_{i=1}^n x_i x_i' \sigma^2 w_i \quad (4)$$

となることを証明した。ここで、White (1980) は(4)を導くのに $\omega$ の値に制約を置いていないことに注意されたい。我々は、興味のあるパラメーター $\omega$ を(4)より抽出するため、以下の多変量線形回帰モデルを考えた。

$$\hat{H} = H + V \quad (5)$$

ここで、 $V$ は $K \times K$ の誤差行列で $\hat{H}$ と $H$ を正値定符号行列にするものである。そして、(5)にvechオペレーター（対称行列の下三角部分を列ベクトル化するオペレーター）を作用させて式を整理すると、 $\sigma^2 w$ をパラメーターとする以下の線形回帰モデルが導ける（この導出は付録を参照のこと）。

$$\hat{h} = \chi_n \sigma^2 \omega + v \quad (6)$$

ここで、 $\hat{h} = \text{vech}(\hat{H})$ は $K' \times 1$ の被説明変数ベクトル、 $\chi_n = [\text{vech}(x_1 x_1'), \dots, \text{vech}(x_n x_n')]$ は $K' \times n$ の説明変数行列、 $v = \text{vech}(V)$ は $K' \times 1$ の誤差ベクトル、 $K' \equiv K(K+1)/2$ である。後は、(6)をOLSで推定すれば、パラメーター $\sigma^2 \omega$ の推定量 $\widehat{\sigma^2 \omega}$ が得られる。ただし、求めているのは $\widehat{\sigma^2 \omega}$ ではなく $\hat{\omega}$ であることに注意が必要である。 $\text{tr}(\Omega) = n$ なる基準化より、 $\hat{\omega}$ は以下のように求められる。

$$\hat{\omega} = \frac{1}{l' \widehat{\sigma^2 \omega} / n} \widehat{\sigma^2 \omega} \quad (7)$$

ここで、 $l = (1, 1, \dots, 1)' \sim n \times 1$ である。つまり、我々は(2)で $\sigma^2$ を推定し、(7)で $w$ を推定することで目的である $\sigma^2 \Omega$ の推定を行う。

ここで、回帰モデル(6)は標本数が $K' \equiv K(K+1)/2$ でパラメーターの数が $n$ であることに注意されたい。通常、回帰モデル(1)において自由度 $n - K$ はかなり大

きいと考えられるので、回帰モデル(6)においては自由度  $K' - n$  がかなり小さく、負になるとも考えられる。次節では、この問題を扱う。

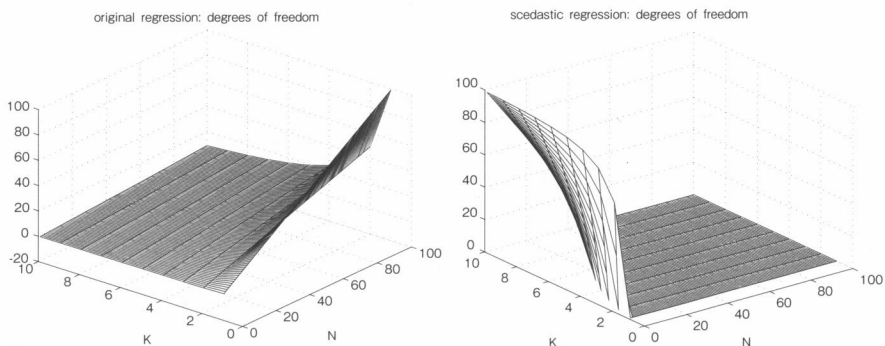
### 3 $\omega$ の識別

#### 3.1 モデルの自由度

前節で、我々は  $\omega$  を推定するために回帰モデル(6)を使ったが、実はこのモデルの自由度はしばしば負になる。このモデルのパラメーターの数は  $n$ 、標本数は  $K'$  なので、例えば  $n = 50$ 、 $K = 3$  とすると、パラメーターの数が50となり標本数  $K' = (3 \times 4)/2 = 6$  を大きく上回ってしまう。つまり、自由度は負になり  $\omega$  は識別されない。(6)において  $\omega$  が識別されるには(6)の自由度  $df_{\omega} = K' - n$  もしくは自由度率  $df_{\omega}\% = 100(K' - n)/K'$  が正である必要がある。また、回帰モデル(6)を推定するには回帰モデル(1)の推定残差が必要なので、もちろん(1)の自由度  $df = n - K$  もしくは自由度率  $df\% = 100(n - K)/n$  も正でなければならない。

図1 回帰モデル(1)の自由度率

図2 回帰モデル(6)の自由度率

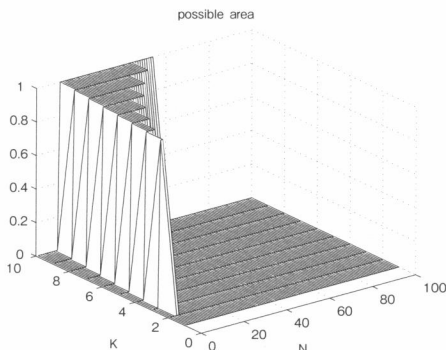


注 図1は  $1 \leq K \leq 10$ 、 $1 \leq n \leq 100$  に対する回帰モデル(1)の自由度率  $df\% = 100(n - K)/n$  を、図2は回帰モデル(6)の自由度率  $df_{\omega}\% = 100(K' - n)/K'$  を表している。自由度率が負の部分は0としている。

これら2つの自由度率がどのような  $n$  と  $K$  の組み合わせに対し正になるかを調べると図1、2のようになる。図1、2は各々回帰モデル(1)、(6)の自由度率を表しており、見やすくするために自由度率が負の部分は0としている。これらの図より、 $df\%$  は  $K$  に比べて  $n$  が大きいときに、 $df_{\omega}\%$  は  $K$  に比べて  $n$  が小さいときに

高くなる傾向にあることが分かる。

図3  $\omega$ が識別可能な $n$ と $K$ の組み合わせ



注 この図は $1 \leq K \leq 10$ 、 $1 \leq n \leq 100$ に対し $\omega$ が識別可能な $n$ と $K$ の組み合わせを示している。識別可能な組み合わせは1、不能な組み合わせは0としている。

$\omega$ を識別するにはこれら2つの自由度率が共に正になる必要があるが、そのような $n$ と $K$ の組み合わせは図3に示されている。図3を見ると、確かに $\omega$ は識別される場合もあるが、識別不能な場合の方が多い。例えば、 $K=6$ の時には、 $\omega$ が識別されるのは $7 \leq n \leq 20$ の時のみである。このように、 $\omega$ は識別性は $n$ と $K$ の危ういバランスの上に成り立っている。

### 3.2 人工的な説明変数を追加したモデルでの識別可能性

前小節で見たように、モデル(1)、(6)を使う限りは $\omega$ を識別できない場合が多くある。特に、回帰モデル(1)の標本数 $n$ が増えると回帰モデル(6)のパラメーターの数が増え、 $\omega$ を識別できなくなるのが問題である。本小節では、この問題を解決する方法を提案する。

我々は、 $n$ が大きいときには回帰モデル(6)の標本数 $K' \equiv K(K+1)/2$ をその $n$ に対応させて増やせば良いと考えた。つまり、回帰モデル(1)に何か説明変数を追加して $K$ を増やせば良い。しかし、恣意的に説明変数を追加したのでは(1)の推定結果が変わってしまい、望ましくない。(1)は正しく特定化されたモデルなので、その推定結果を変えずに説明変数を追加しなければならないのである。

そこで、我々は回帰モデル(1)に、

$$W[y, X] = 0 \quad (8)$$

なる性質を持つ  $n \times K_w$  の説明変数行列  $W$  を追加した。つまり、(1)の代わりに次のような回帰モデルを推定する。

$$y = Z\gamma + \varepsilon \quad (9)$$

ここで、 $Z = [X, W]$  は  $n \times (K + K_w)$  の説明変数行、 $\gamma = [\beta, \delta']$  は  $(K + K_w) \times 1$  のパラメーターベクトル、 $\delta$  は  $K_w \times 1$  のパラメーターベクトルである。(9)における  $\beta$  の OLS 推定量は(1)のそれと同じであることが容易に導出できる(付録を参照のこと)。具体的には、

$$\hat{\gamma} = \begin{bmatrix} \hat{\beta} \\ 0 \end{bmatrix} \quad (10)$$

ここで、 $\hat{\gamma}$  は  $\gamma$  の OLS 推定量である。(10)は、(1)から得る推定残差と(9)からの推定残差も同じであることも示している。つまり、(8)なる性質を持つ説明変数を(1)に追加すれば、推定結果を全く変えずに  $K$  を大きくできる。後は、 $n$  に応じて  $K_w$  を適切に大きくして(9)を推定して  $\hat{h}$  と  $\chi_n$  を得れば、回帰モデル(6)における標本数  $(K + K_w)(K + K_w + 1)/2$  がパラメーターの数  $n$  を上回り  $\omega$  の識別が可能となる。例えば、 $n = 50$ 、 $K = 3$  の時には  $K_w = 7$  個以上の説明変数  $W$  を追加すれば良い。ただし、47 個以上追加すると今度は(9)が識別不能になるので注意が必要である。

ここで、 $W$  の計算法について述べておく。一般に、 $a \times b$  ( $a > b$  とする) の行列  $A$  には、 $A_{\perp} A = 0$  なる性質を持つ  $a \times (a - b)$  行列  $A_{\perp}$  が存在することが知られている。 $A_{\perp}$  は、例えば行列言語 GAUSS の “null” コマンドを使って簡単に計算できる。このコマンドを使えば  $y_{\perp}$  や  $X_{\perp}$  が簡単に計算でき、(8)なる性質を持つ  $W$  も計算できる。 $W$  の計算には、元のデータ  $y$ 、 $X$  以外に追加的なデータは必要ないことに注意されたい。

#### 4 モンテカルロ実験

本節では、モンテカルロ実験を行って2、3節で提案された不均一分散の推定法が適切に機能しているか確かめる。実験を行う前に、推定の手順を整理しておく。

Step 1:  $n$ と $K'$ のどちらが大きい確認する。 $n < K'$ ならば回帰モデル(1)をOLSで推定して推定残差 $e$ と $\hat{h}$ 、 $\chi_n$ を計算する。 $n > K'$ ならば人工的な説明変数 $W \sim n \times K_w$ を $n < (K + K_w)(K + K_w + 1)/2$ となるように作り、回帰モデル(9)をOLSで推定して $e$ と $\hat{h}$ 、 $\chi_n$ を計算する。

Step 2: Step 1で計算した $e$ を(2)に代入し、 $\sigma^2$ の推定量 $\hat{\sigma}^2$ を計算する。Step 1で計算した $\hat{h}$ と $\chi_n$ を使って回帰モデル(6)をOLSで推定し、(7)より $\omega$ の推定量 $\hat{\omega}$ を計算する。

Step 3: Step 2で得た $\hat{\sigma}^2$ と $\hat{\omega}$ を掛け合わせて、不均一分散の推定量 $\hat{\sigma}^2 \hat{\omega}$ を得る。

モンテカルロ実験には、次のデータ生成過程を使った。

$$y = X\beta + \varepsilon, \varepsilon \sim N(0, \sigma^2 \Omega), X = (1, x_2, \dots, x_K), x_j \sim N(\mu_j, \lambda_j I_n), j = 2, \dots, K$$

ここで、 $\varepsilon$ と $X$ は独立、 $x_j$ も各々独立である。標本数や説明変数の数、パラメータの値は次のように設定した。

ケース1 ( $n < K'$ のケース)

$n = 25$ 、 $K = 7$ 、 $\beta = (1, 3, -2, 1, -3, 0.5, -0.6)'$ 、 $(\mu_2, \dots, \mu_7) = (3, 6, 0, 2, 5, 1)$ 、 $(\lambda_2, \dots, \lambda_7) = (1, 2, 1, 1, 0.5, 2)$ 、 $\sigma^2$ と $\omega$ は表1参照のこと。

ケース2 ( $n > K'$ のケース)

$n = 50$ 、 $K = 3$ 、 $\beta = (1, 3, -2)'$ 、 $(\mu_2, \mu_3) = (3, 6)$ 、 $(\lambda_2, \lambda_3) = (1, 4)$ 、 $\sigma^2$ と $\omega$ は表2参照のこと。

3節で議論したように、我々の推定量は $n < K'$ の場合と $n > K'$ の場合で手順が異なるので、実験では上のように両方の場合を考慮した。ケース1は $\omega$ が識別できる場合である。それに対し、ケース2はそのままでは $\omega$ が識別できないので3.2節で提案した人工的な説明変数 $W$ を加えて推定を行う場合である。この実験では、7個の人工的な説明変数を加えた。



表1 実験結果(ケース1)

index	true	mean	s.d.
1	1.31	1.10	1.55
2	2.60	2.41	3.43
3	0.24	0.38	0.54
4	0.62	0.72	1.03
5	0.28	0.63	0.89
6	1.27	1.32	1.84
7	3.66	3.35	4.73
8	1.76	1.77	2.52
9	0.82	1.00	1.41
10	0.16	0.29	0.41
11	3.14	2.32	3.28
12	0.41	0.72	1.01
13	0.87	0.87	1.23
14	1.73	1.61	2.27
15	0.11	0.26	0.37
16	0.99	0.92	1.30
17	0.72	0.81	1.15
18	0.57	0.79	1.12
19	3.06	2.31	3.28
20	1.37	1.12	1.57
21	0.46	0.60	0.86
22	1.31	1.26	1.79
23	0.45	0.72	1.01
24	0.09	0.31	0.44
25	0.42	0.62	0.88

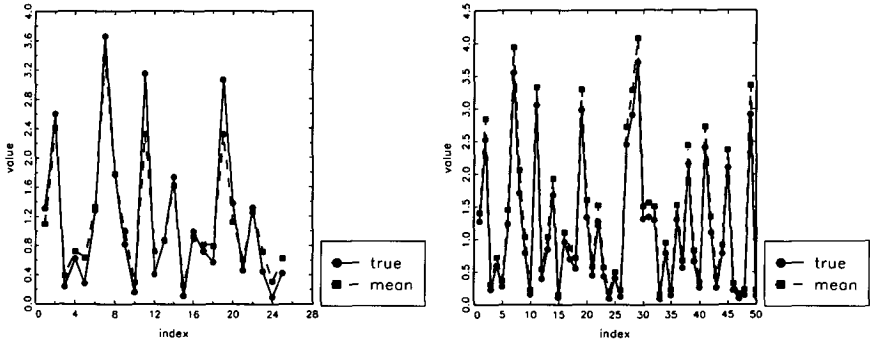
表2 実験結果(ケース2)

index	true	mean	s.d.	index	true	mean	s.d.
1	1.27	1.40	1.96	26	0.12	0.22	0.32
2	2.53	2.85	3.97	27	2.45	2.72	3.83
3	0.23	0.31	0.44	28	2.90	3.28	4.57
4	0.60	0.72	1.02	29	3.71	4.08	5.61
5	0.27	0.36	0.52	30	1.32	1.51	2.11
6	1.24	1.45	2.01	31	1.34	1.56	2.19
7	3.55	3.94	5.66	32	1.29	1.50	2.10
8	1.71	2.06	2.85	33	0.09	0.15	0.22
9	0.79	1.04	1.51	34	0.79	0.95	1.33
10	0.16	0.23	0.32	35	0.14	0.23	0.33
11	3.05	3.33	4.64	36	1.30	1.52	2.15
12	0.39	0.54	0.76	37	0.56	0.67	0.95
13	0.85	1.04	1.45	38	2.16	2.44	3.37
14	1.68	1.93	2.76	39	0.66	0.83	1.16
15	0.11	0.15	0.21	40	0.26	0.35	0.49
16	0.96	1.11	1.57	41	2.41	2.72	3.79
17	0.70	0.87	1.24	42	1.11	1.35	1.89
18	0.55	0.71	1.02	43	0.26	0.43	0.61
19	2.97	3.29	4.50	44	0.78	0.91	1.27
20	1.33	1.60	2.27	45	2.10	2.37	3.34
21	0.45	0.58	0.81	46	0.22	0.32	0.46
22	1.28	1.52	2.13	47	0.09	0.17	0.24
23	0.43	0.57	0.82	48	0.14	0.23	0.33
24	0.09	0.20	0.29	49	2.92	3.36	4.69
25	0.41	0.50	0.71	50	0.13	0.22	0.31

注：これら2つの表において、“index”は観測値の番号、“true”は真の $\sigma^2\omega$ の値、“mean”は10000個の $\hat{\sigma}^2\hat{\omega}$ の平均、“s.d.”は10000個の $\hat{\sigma}^2\hat{\omega}$ の標準偏差を示す。

以上の設定で10000回の繰り返しのモンテカルロ実験を行ったところ、ケース1では表1の結果を得た。この表より、10000個の推定値の平均は真の値とほぼ一致していることが解る。例えば、 $i=1$ では真の値が1.31であるのに対し推定値の平均は1.10であり、 $i=2$ では真の値2.60に対し推定値の平均は2.41である。この結果は、この推定量が適切に機能することを示している。推定値の平均と真の値がほぼ一致していることは、図4を見ればより明らかであろう。一方で、表1は、この推定量の標準誤差が大きいことも示唆している。例えば、 $i=1$ では10000個の推定値の標本標準偏差が1.55である。標準誤差が大きいことは、我々の推定量の1つの問題である。

図4 推定値の平均と真の値 (ケース1) 図5 推定値の平均と真の値 (ケース2)



注：これら2つの図において、“true”は真の $\sigma^2 \omega$ の値、“mean”は10000個の $\hat{\sigma}^2 \hat{\omega}$ の平均を表す。

ケース2の結果は表2に示されているが、ケース1の結果とほぼ同じである。つまり、推定値の平均は真の値とほぼ一致した(図5も参照のこと)。この結果より、3.2節で提案した $\omega$ の識別性の問題を回避する方法が適切に機能していると考えられる。しかし、ケース1と同様に、標準偏差が大きという問題もある。

## 5 まとめと今後の展望

本稿では、線形回帰モデルにおいて、不均一な誤差項分散の推定法を提案した。我々は、White (1980) のHCEにvechオペレーターを作用させることで不均一分散をパラメーターとする線形回帰モデルを導き、それをOLSで推定することで不均一分散の推定を行った。このとき、推定したいパラメーターである不均一分散は多くの場合識別できないという問題が生じるが、本稿ではある人工的な説明変数を使うことでこの問題を解決した。モンテカルロ実験により、この推定法は若干の問題もあるものの適切に機能することが確かめられた。この推定法は不均一分散に一切の構造を仮定しないという特徴を持っており、不均一分散の構造に対して頑健なGLS推定やファイナンスの分野でのボラティリティの頑健な推定に応用できると考えられる。

今後の課題としては、この推定法のファイナンスやGLS推定等への応用と、推定量の標準誤差が大きという問題の克服の2点が挙げられる。我々はこの推

定法を、企業毎の株価データを使った横断面方向に異なるボラティリティの推定に応用しようと考えている。このとき、高頻度データを使って企業毎の株価データを時系列方向に集めていわゆるパネルデータを作成し、それを利用して推定量の標準誤差を下げようと試みている。また、近年、不均一分散を持つ回帰モデルにおいて回帰係数に対する推量のパフォーマンスを改善するためにいくつかのブートストラップ法が提案されている (Flachaire (2005) 等参照のこと)。そこで、それらのブートストラップ法が我々の推定量のパフォーマンスの改善、つまり標準誤差の減少にも使えるか検討している。

#### 参考文献

- Flachaire, E. (2005) : “Bootstrapping heteroskedastic regression models : wild bootstrap vs. pairwise bootstrap,” *Computational Statistics & Data Analysis*, 49, 361–376.
- Greene, W.H. (2003) : *Econometric Analysis*, 5th ed., Prentice-Hall.
- Lütkepohl, H. (1996) : *Handbook of Matrices*, John Wiley & Sons.
- White, H. (1980) : “A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity,” *Econometrica*, 48, 817–838.

#### 付録

(4)の導出：ここでは、本文(4)の導出、つまり  $h \equiv \text{vech}(H) = \chi_n \omega$  の導出を行う。そのために必要ないくつかの補題を以下に示す。

補題 1 :  $\text{vec}(H) = (X' \otimes X') \text{vec}(\Omega) \sim K^2 \times 1$

証明 :  $\text{vec}(ABC) = (C' \otimes A) \text{vec}(B)$  なる関係式より得られる。この関係式については Lütkepohl (1996, p.97) 等を参照のこと。

補題 2 :  $X' \otimes X' = [(x_1 \otimes x_1), \dots, (x_1 \otimes x_n)], \dots, [(x_n \otimes x_1), \dots, (x_n \otimes x_n)] \sim K^2 \times n^2$

証明 :  $X' = (x_1, \dots, x_n) \sim K \times n$  であることより上記の結果を得る。

補題 3 :  $\text{vec}(H) = [(x_1 \otimes x_1), \dots, (x_n \otimes x_n)] \omega$

証明 :  $\text{vec}(\Omega) = \text{vec}(\text{diag}(\omega)) = (\omega_1, 0, \dots, 0; 0, \omega_2, 0, \dots, 0; \dots; \omega_n) \sim n^2 \times 1$  である。

ここで、 $\text{diag}(\omega)$  は  $\omega$  を対角要素に持つ対角行列を示す。これより、 $n^2 \times 1$  ベクトルの  $\text{vec}(\Omega)$  が  $X' \otimes X'$  から  $(x_i \otimes x_i)$ ,  $i = 1, \dots, n$  のみを抜き取る働きがあることが分かる。このことと補題 1、2 を合わせて補題 3 の結果を得る。

補題4 :  $x_i \otimes x_i = \text{vec}(x_i x_i') \sim K^2 \times 1$

証明 :  $\text{vec}(AB) = (B' \otimes A) \text{vec}(I_p)$ なる関係式より得られる。この関係式については Lütekepohl (1996, p.97) 等参照のこと。

ここで、 $\text{vec}(\cdot)$  を  $\text{vech}(\cdot)$  に変換する  $L_{K'}$  を導入する。

定義 :  $L_{K'}$  を次のように定義する。

$$L_{K'} = \begin{bmatrix} I_K & & & & & & 0 \\ & 0 & & & & & \\ & \vdots & I_{K-1} & & & & \\ & & \ddots & I_{K-2} & & & \\ & & & \ddots & \dots & & \\ & 0 & & & & \dots & \\ & & & & & & 0 \cdots 0 \ 1 \end{bmatrix} \sim K' \times K^2$$

$L_{K'}$  は、その対角にある部分行列が  $I_K$  から  $I_{K-1}$ 、 $I_{K-2}$ 、と縮み、最後は1になるという構造を持っている。縮んだ部分は0で置き換えられる。

この  $L_{K'}$  は、次の補題に示すように、 $\text{vec}(\cdot)$  を  $\text{vech}(\cdot)$  に変換する働きを持つ。

補題5 :  $K \times K$  行列  $A$  に対し、 $\text{vech}(A) = L_{K'} \text{vec}(A) \sim K' \times 1$  となる。

証明 :  $A = (a_1, \dots, a_K) \sim K \times K$  とし、 $a_j \sim K \times 1$  とする。そして、

$$a_{j(-1,2)} = (a_{3j}, a_{4j}, \dots, a_{Kj})' \sim (K-2) \times 1$$

のように書く。つまり、 $a_{j(-1,2)}$  は  $a_j$  の最初の2つの要素を除いたものである。これより、

$$L_{K'} \text{vec}(A) = L_{K'} \begin{pmatrix} a_1 \\ \vdots \\ a_K \end{pmatrix} = L_{K'} \begin{pmatrix} a_1 \\ \dots \\ 0 \\ a_{2(-1)} \\ \dots \\ 0 \\ 0 \\ a_{3(-1,2)} \\ \dots \\ \vdots \\ a_{KK} \end{pmatrix} = \text{vech}(A)$$

を得る。

補題6 :  $\text{vech}(x_i x_i') = L_{K'} \text{vec}(x_i x_i')$

証明 : 省略。

では、これらの補題を使って  $\text{vech}(H) = \chi_n \omega$  を導出する。補題2、3、4の結果を

補題1の関係式に適用すると

$\text{vec}(H) = [\text{vec}(x_1 x_1'), \dots, \text{vec}(x_n x_n')] \omega$   
 を得る。さらに、補題5、6より以下を得る。

$$\begin{aligned} \text{vech}(H) &= L_{K'} \text{vec}(H) \omega \\ &= [L_{K'} \text{vec}(x_1 x_1'), \dots, L_{K'} \text{vec}(x_n x_n')] \omega \\ &= \chi_n \omega \end{aligned}$$

ここで、 $\chi_n = [\text{vech}(x_1 x_1'), \dots, \text{vech}(x_n x_n')] \sim K' \times n$ である。以上で本文(4)が導出できた。

(10)の導出:ここでは、本文(10)の導出を行う。

$$\begin{aligned} \hat{y} &\equiv \left\{ \begin{bmatrix} X' \\ W' \end{bmatrix} \begin{bmatrix} X & W \end{bmatrix} \right\}^{-1} \begin{bmatrix} X' \\ W' \end{bmatrix} y \\ &= \begin{bmatrix} X'X & 0 \\ 0 & W'W \end{bmatrix}^{-1} \begin{bmatrix} X'y \\ 0 \end{bmatrix} = \begin{bmatrix} \hat{\beta} \\ 0 \end{bmatrix} \end{aligned}$$