응용통계연구 (2008)
21(4), 561–569

재무 시계열 자료 분석을 위한 로버스트 추정방법

김삼용[1]

[1]중앙대학교 통계학과

(2008년 4월 접수, 2008년 5월 채택)

요 약

본 논문은 재무 시계열 자료에서 흔히 나타나는 이상치를 처리하기 위하여 이중 로버스트 추정함수를 제시하였다. 이중 로버스트 추정 방정식의 해인 로버스트 추정치를 이용하여 ARCH모형과 GARCH 모형 하에서 이상치를 처리하였다. 또한 실제 주가자료를 응용하여 기존의 최소제곱추정치보다 로버스트 추정치나 이중 로버스트 추정치의 성능이 우수함을 보였다.

주요용어: 로버스트 추정, GARCH 모형, 주가지수.

1. 서론

환율이나 이자율과 같은 금융시계열 자료는 시간의 변화에 따라 변동성이 매우 급변하는 특징이 있다. 이러한 변동성은 Box와 Jenkins (1976)가 제안한 선형시계열 모형인 ARIMA 모형으로는 제대로 잡아 낼 수가 없다고 알려져 있다. 이러한 변동성을 설명하기 위하여 Engle (1982)은 처음으로 조건부 이분산성 모형인 ARCH 모형을 제안하였고 Bollerslev (1986)은 일반화 ARCH(Generlaized ARCH: GARCH) 모형을 제시하였다. 한편으로 ARCH 모형이나 GARCH 모형 하에서 추정된 잔차의 형태를 살펴보면 첨도의 값이 과다하게 큰 것을 알 수 있으며 이것은 이상치가 존재하는 경우 GARCH 모형으로는 제대로 변동성을 설명할 수 없는 상황이 발생한다는 것을 의미한다. 자료에 있어서 이상치의 존재는 모형의 적합성과 모수 추정치에 많은 영향을 주고 있으며 특히 최소제곱추정치나 정규 분포와의 최대우도 추정치의 정확도를 떨어뜨리는 역할을 하고 있다. 이상치를 다루는 문제는 많은 연구의 의해서 진행되어 왔는데 Chan과 Cheung (1994)는 Generalized-M 추정방법으로 threshold AR 모형에서의 이상치를 다루었다. Franses와 Ghijsels (1999)는 GARCH 모형 하에서 가법적 이상치(additive outliers)를 탐지하고 처리하는 방법을 제시하였다. 최근에 Charles와 Darne (2005)은 이를 발전시켜 GARCH 모형 하에서 innovative 이상치를 다루는 방법을 제시하였다. 한편 Godambe (1985)는 선형 시계열 모형의 모수 추정을 위해 추정함수 방법을 제시 하였다. 그리고 Kim과 Hwang (2007)은 이중 로버스트 추정함수를 이용하여 AR(1)-ARCH(1) 모형의 모수를 추정하고 추정치의 점근적인 성질을 규명하였다. 본 연구에서는 Kim과 Hwang (2007)이 제시한 이중 로버스트 추정함수를 응용하여 ARCH 모형 및 GARCH 모형 하에서 이상치가 존재하는 경우 이를 처리하는 방법을 실제 주가자료를 응용하여 그 효율성을 보이고자 한다.

본 연구는 2007년도 중앙대학교 학술연구비 지원에 의한 것임.
[1](156-756) 서울시 동작구 흑석동, 중앙대학교 통계학과, 부교수. E-mail: sahm@cau.ac.kr

2. GARCH(1,1) 모형과 로버스트 추정함수

로버스트 추정함수를 위해 먼저 다음과 같은 AR(1)-GARCH(1,1) 모형을 고려하기로 한다.

$$
\begin{aligned}
y_t &= \phi y_{t-1} + \epsilon_t, \\
\epsilon_t &= \sqrt{h_t} e_t, \\
h_t &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}, \quad t = 1, 2, ..., n, \\
e_t &\sim iid(0, \sigma^2), \quad \alpha_0,\ \alpha_1 > 0, \quad \beta_1 > 0, \quad \alpha_1 + \beta_1 < 1,
\end{aligned}
$$

$h_t = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1}$에서 $\beta_1 = 0$이면 AR(1)-ARCH(1,1) 모형이 된다.

GARCH 모형은 재무 자료의 변동성을 기존의 ARIMA 모형 보다는 효율적으로 탐지 할 수 있으나 이상치의 존재 시 추정치에 약점이 있음을 알 수 있다 (Franses와 Ghijsels, 1999). 여기서 먼저 Kim과 Hwang (2007)이 제안한 이중 로버스트 함수를 근거로 이 방정식의 해인 로버스트 추정치를 구하고 기존의 최소제곱추정치 및 로버스트 추정치와 비교하기로 한다.

다음의 이중 로버스트 추정함수를 생각해 보자.

$$S_n(\theta) = \sum_{t=1}^{n} \Psi_1 \left(\frac{\epsilon_t(\theta)}{\sqrt{h_t}} \right) \Psi_2 \left(\frac{d}{d\theta} \frac{\epsilon_t(\theta)}{\sqrt{h_t}} \right),$$

여기서 $\theta = (\phi,\ \alpha_0,\ \alpha_1,\ \beta_1$이고 Ψ_1과 Ψ_2는 위계 함수이자 θ에 대하여 미분 가능 함수이다. 일반적으로 Ψ_1과 Ψ_2는 Huber의 함수 (Huber, 1981)를 사용하며 만일 $\Psi_2(x) = x$라면 기존의 로버스트 함수와 같게 된다 (Denby와 Martin, 1979). Kim과 Hwang (2007)이 제안한 이중 로버스트 함수를 이용하여 적절한 조건하에서 GARCH 모형으로 확장하여 다음의 정리를 보일 수 있다.

정리 2.1 적절한 조건하에서

(1) $\widehat{\theta_n}$은 θ로 확률 수렴한다.

(2) $\sqrt{n}((\widehat{\theta_n} - \theta)$는 $N(0, (A(\theta)B^{-1}A(\theta)^T)^{-1})$로 분포 수렴한다. 여기서 $\widehat{\theta_n}$은 이중 로버스트 방정식의 해이고 $A(\theta) = E(d/d\theta s_n(\theta))$이며 $B(\theta) = E(S_n(\theta)S_n(\theta)^T)$이다.

증명: Kim과 Hwang (2007) 에서는 AR(1)-ARCH(1) 모형에 대하여 증명하였고 이것을 적절히 확장하여 AR(1)-GARCH(1,1) 모형에도 적용할 수 있으므로 증명은 생략하기로 한다. □

3. 주식자료 분석을 통한 성능 비교

본 논문에서는 5년간(1997.1.03 ∼ 2001.12.28) KOSPI 지수와 한국의 4개사(한미약품, LG, 대우증권, 현대모비스)의 상장된 주식 가격의 일별 자료를 이용하여 로그-수익률(log-return) 즉, y_t를 예측하고자 한다. 로그 수익률은 다음과 같으며 $y_t = \log(x_t/x_{t-1}) \times 100$ x_t는 t시점에서의 주가이다.

예측에 사용된 모형으로는 AR(1)-ARCH(1) 모형과 AR(1)-GARCH(1,1) 모형이며 각 모형에 대하여 최소제곱추정치와 기존의 로버스트 추정치 그리고 이중 로버스트 추정치를 구하고, 예측값의 RMSE를 산출하여 비교하였다. 여기서

$$\text{RMSE} = \sqrt{\sum_{t=1}^{n} (y_t - \widehat{y_t})^2}$$

이다.

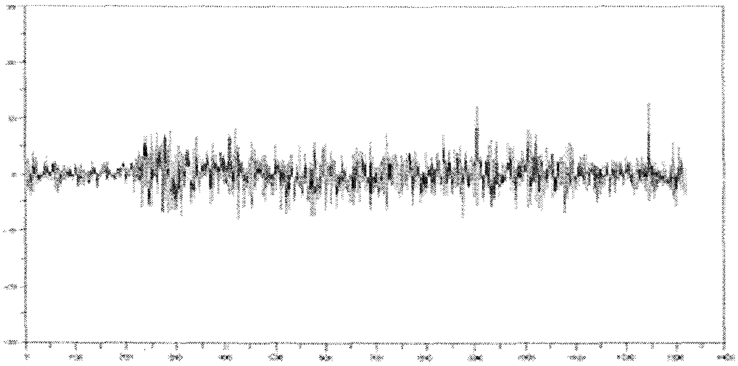

그림 **3.1.** KOSPI의 시계열 시도표

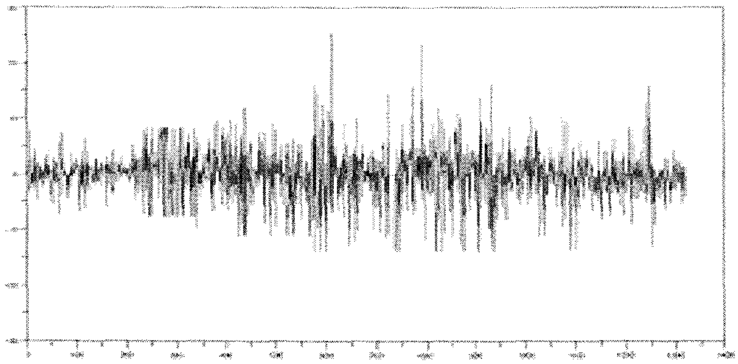

그림 **3.2.** 현대모비스의 시계열 시도표

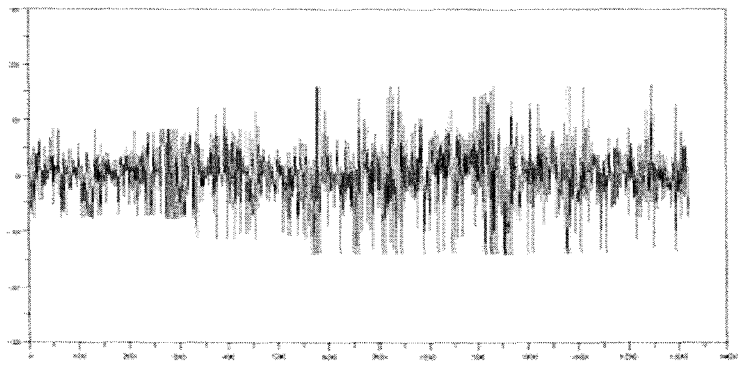

그림 **3.3.** 대우증권의 시계열 시도표

먼저 그림 3.1~3.4는 KOSPI 지수 및 개별 주가의 수익률에 관한 시도표이고, 그림 3.
plot이다.

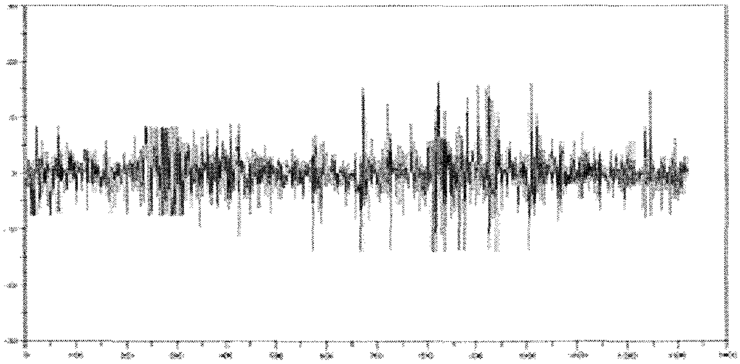

그림 3.4. 한미약품의 시계열 시도표

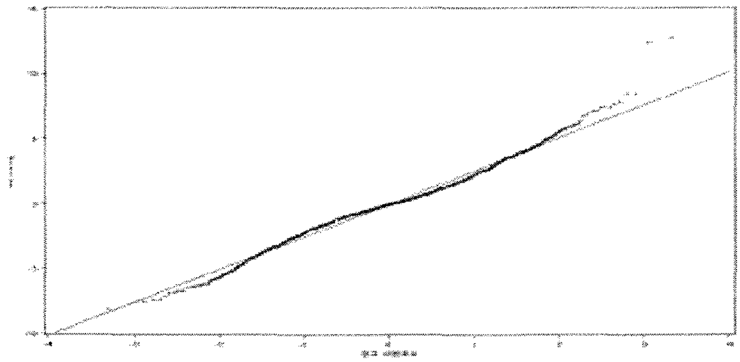

그림 3.5. KOSPI의 Q-Q PLOT

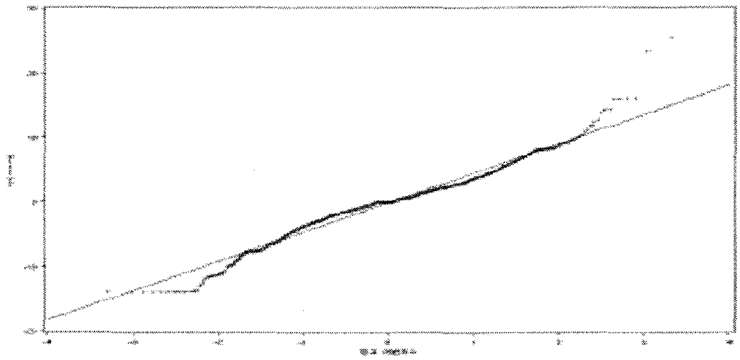

그림 3.6. 현대모비스의 Q-Q PLOT

위의 자료를 보면 개별 주가의 변동성이 KOSPI 주가 보다 큰 것을 알 수 있다. 이것은 개별 자료. 동성이 합산되어 KOSPI 주가의 변동을 순화(smoothing) 시킴을 의미한다.

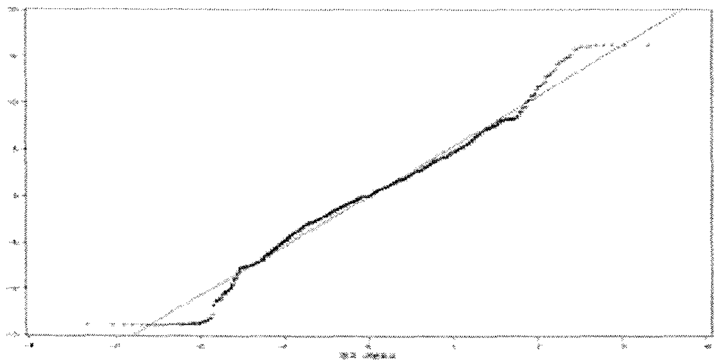

그림 **3.7.** 대우증권의 Q-Q PLOT

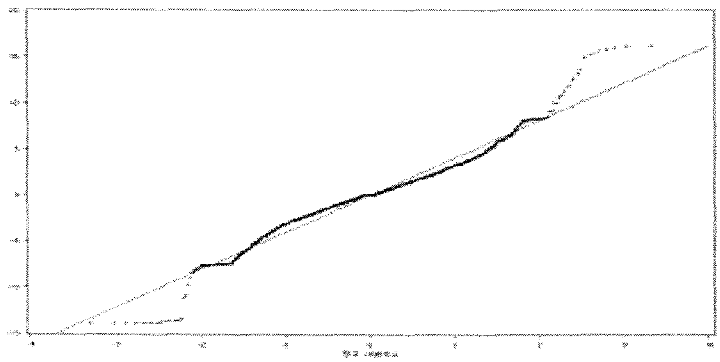

그림 **3.8.** 한미약품의 Q-Q PLOT

표 **3.1.** SHAPIRO-WILK TEST

종목	통계량	p-value
KOSPI	0.9816	<0.0001
현대모비스	0.9645	<0.0001
대우증권	0.9812	<0.0001
한미약품	0.9590	<0.0001

또한 Shapiro-Wilk test의 결과 (표 3.1)를 보면 정규성을 만족한다고 볼 수 없을 뿐만 아니라 자료가 IMF시기를 포함하고 있어 변동성이 심하고 특히 이상치가 많이 발생함을 볼 수 있다. 이러한 연유로 이상치에 민감한 최소제곱추정치나 기존의 로버스트 추정치보다 제안된 이중 로버스트 추정치의 활용이 요구된다 할 수 있다. 기존의 로버스트 추정치보다 이중 로버스트 추정치를 제안하는 이유는 시계열 자료가 기본적으로 종속적인 자료이고 이러한 상황에서 오차 및 과거자료에 동시에 변동성을 효과적으로 제어 할 수 있기 때문이다. 아래 표 3.2~3.13는 5년간(1997.1.03 ~ 2001.12.28) KOSPI 지수와 한국의 3개사(한미약품, 대우증권, 현대모비스)의 종가를 이용하여 로그 수익률 자료를 만들어(1319개) 1000개의 자료를 이용하여 모수를 추정하였고, 나머지 319개의 자료를 이용하여 RMSE를 산출 하였다. 모수

표 **3.2.** $k = 0.5$, AR(1)-ARCH(1) 모형의 모수 추정치

	Least Square			Robust			Double Robust		
	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$
KOSPI	0.1387	0.0058	0.1059	0.1387	0.0058	0.1060	0.2222	0.0058	0.1049
현대모비스	0.0548	0.0204	0.0975	0.0549	0.0204	0.0975	0.0789	0.0204	0.0905
대우증권	0.0867	0.0244	0.1952	0.0868	0.0244	0.1953	0.0521	0.0245	0.1572
한미약품	0.0581	0.0128	0.2966	0.0582	0.0128	0.2968	0.0179	0.0128	0.2980

표 **3.3.** $k = 0.5$, AR(1)-ARCH(1) 모형의 RMSE

	Least Squares	Robust	Double Robust
KOSPI	2.4081	2.4080	2.4456
현대모비스	3.8073	3.8073	3.8052
대우증권	5.0395	5.0396	5.0230
한미약품	3.3176	3.3176	3.3176

표 **3.4.** $k = 1$, AR(1)-ARCH(1) 모형의 모수 추정치

	Least Square			Robust			Double Robust		
	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$
KOSPI	0.1387	0.0058	0.1059	0.1386	0.0058	0.1060	0.4740	0.0450	0.0082
현대모비스	0.0548	0.0204	0.0975	0.0550	0.0204	0.0975	0.0329	0.0205	0.1086
대우증권	0.0867	0.0244	0.1952	0.0867	0.0244	0.1953	0.0238	0.0243	0.1906
한미약품	0.0581	0.0128	0.2966	0.0581	0.0128	0.2968	0.0672	0.0125	0.2978

표 **3.5.** $k = 1$, AR(1)-ARCH(1) 모형의 RMSE

	Least Squares	Robust	Double Robust
KOSPI	2.4081	2.4080	2.6474
현대모비스	3.8073	3.8073	3.8111
대우증권	5.0395	5.0395	5.0139
한미약품	3.3176	3.3176	3.3197

표 **3.6.** $k = 1.5$, AR(1)-ARCH(1) 모형의 모수 추정치

	Least Square			Robust			Double Robust		
	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$
KOSPI	0.1387	0.0058	0.1059	0.1385	0.0058	0.1060	0.1539	0.1050	0.1276
현대모비스	0.0548	0.0204	0.0975	0.0548	0.0204	0.0975	0.0608	0.0205	0.1051
대우증권	0.0867	0.0244	0.1952	0.0867	0.0244	0.1953	0.0285	0.0240	0.2150
한미약품	0.0581	0.0128	0.2966	0.0579	0.0128	0.2968	0.0084	0.0128	0.3123

의 추정치는 다음과 같은 one-step 추정방정식의 해를 사용하였다.

$$\widehat{\theta_n} = \widehat{\theta_{LS}} - \left(\frac{S_n'\left(\widehat{\theta_{LS}}\right)}{n} \right)^{-1} \left(\frac{S_n\left(\widehat{\theta_{LS}}\right)}{n} \right),$$

여기서는 $\widehat{\theta_{LS}}$는 모수 추정 시 초기치(intial value)로 사용하였다.

표 3.7. $k = 1.5$, AR(1)-ARCH(1) 모형의 RMSE

	Least Squares	Robust	Double Robust
KOSPI	2.4081	2.4080	2.4137
현대모비스	3.8073	3.8073	3.8066
대우증권	5.0395	5.0395	5.0151
한미약품	3.3176	3.3175	3.3104

표 3.8. $k = 0.5$, AR(1)-GARCH(1,1) 모형의 모수 추정치

	Least Square				Robust				Double Robust			
	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\beta_1}$	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\beta_1}$	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\beta_1}$
KOSPI	0.1381	0.0196	0.0808	0.9208	0.1381	0.0196	0.0808	0.9209	0.1385	0.0288	0.0808	0.9207
현대모비스	0.0308	0.1321	0.0799	0.9184	0.0307	0.1323	0.0799	0.9185	−0.0710	0.0796	0.0799	0.9020
대우증권	0.0722	0.9008	0.1082	0.8626	0.0721	0.9017	0.1082	0.8628	0.0972	0.5181	0.1082	0.8500
한미약품	0.0374	1.3908	0.1946	0.7296	0.0373	1.3919	0.1946	0.7298	0.0418	1.2572	0.1946	0.7985

표 3.9. $k = 0.5$, AR(1)-GARCH(1,1) 모형의 RMSE

	Least Squares	Robust	Double Robust
KOSPI	3.4629	2.4078	2.4080
현대모비스	5.7063	3.8115	3.8542
대우증권	7.2091	5.0319	5.0457
한미약품	4.7601	3.3138	3.3145

표 3.10. $k = 1$, AR(1)-GARCH(1,1) 모형의 모수 추정치

	Least Square				Robust				Double Robust			
	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\beta_1}$	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\beta_1}$	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\beta_1}$
KOSPI	0.1381	0.0196	0.0808	0.9208	0.1381	0.0196	0.0808	0.9209	0.1441	0.0168	0.0808	0.9208
현대모비스	0.0308	0.1321	0.0799	0.9184	0.0308	0.1323	0.0799	0.9185	0.0059	0.1372	0.0799	0.9185
대우증권	0.0722	0.9008	0.1082	0.8626	0.0722	0.9017	0.1082	0.8628	0.1083	0.1773	0.1082	0.8664
한미약품	0.0374	1.3908	0.1946	0.7296	0.0374	1.3919	0.1946	0.7298	0.0325	1.5467	0.1946	0.7343

표 3.11. $k = 1$, AR(1)-GARCH(1,1) 모형의 RMSE

	Least Squares	Robust	Double Robust
KOSPI	3.4629	2.4078	2.4100
현대모비스	5.7063	3.8115	3.8182
대우증권	7.2091	5.0319	5.0529
한미약품	4.7601	3.3138	3.3131

표 3.12. $k = 1.5$, AR(1)-GARCH(1,1) 모형의 모수 추정치

	Least Square				Robust				Double Robust			
	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\beta_1}$	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\beta_1}$	$\widehat{\phi_1}$	$\widehat{\alpha_0}$	$\widehat{\alpha_1}$	$\widehat{\beta_1}$
KOSPI	0.1381	0.0196	0.0808	0.9208	0.1381	0.0196	0.0808	0.9209	0.1360	0.0106	0.0808	0.9203
현대모비스	0.0308	0.1321	0.0799	0.9184	0.0308	0.1323	0.0799	0.9185	0.0423	0.1370	0.0799	0.9181
대우증권	0.0722	0.9008	0.1082	0.8626	0.0722	0.9017	0.1082	0.8628	0.0800	0.2375	0.1082	0.8650
한미약품	0.0374	1.3908	0.1946	0.7296	0.0374	1.3919	0.1946	0.7298	0.0348	1.4654	0.1946	0.7292

표 3.13. $k = 1.5$, AR(1)-GARCH(1,1) 모형의 RMSE

	Least Squares	Robust	Double Robust
KOSPI	0.7292	2.4078	2.4071
현대모비스	5.7063	3.8115	3.8092
대우증권	7.2091	5.0319	5.0359
한미약품	4.7601	3.3138	3.3134

4. 결론

일반적으로 이상치가 존재하는 경우 일반적인 추정치보다 로버스트 추정치가 더 적합한 것으로 알려져 있다. 본 연구에서 고려한 AR(1)-ARCH(1) 모형과 AR(1)-GARCH(1,1) 모형의 RMSE를 살펴보면 각각의 모형에서 로버스트 추정치나 이중 로버스트 추정치가 최소제곱추정치보다 성능이 더 우수함을 알 수 있다. k값이 0.5이거나 1인 경우에는 로버스트 추정치가 이중 로버스트 추정치보다 차이가 크지 않았거나 오히려 약간은 좋았으나, k값이 1.5인 경우에는 두 모형의 모수에 대한 로버스트 추정치보다 이중 로버스트 추정치가 더 적합을 잘 시키는 것으로 나타났다. 이러한 결과가 나타난 원인으로는 IMF에 의해 주가가 급변하였고, 이에 이상치가 발생한 결과로 사료된다. 따라서 이상치가 존재하는 자료에서는 로버스트나 이중 로버스트 모형을 사용하는 것이 일반적인 모형을 사용하는 것보다 적절하며, 이상치에 민감하게 반응을 해야 하는 경우라면 기존의 로버스트 모형보다 이중 로버스트 모형을 사용하는 것이 적절한 것으로 생각된다. 좀 더 정교하고 복잡한 모형에 이중 로버스트 방법을 적합하는 것도 향후의 연구 방향이 될 수 있을 것이다.

참고문헌

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307–327.

Box, G. E. P. and Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control*, Holden Day, Oakland.

Chan, W. S. and Cheung, S. H. (1994). On robust estimation of the threshold autoregressions, *Journal of Forecasting*, **13**, 37–49.

Charles, A. and Darne, O. (2005). Outliers and GARCH models in financial data, *Economics Letters*, **86**, 347–352.

Denby, L. and Martin, R. D. (1979). Robust estimation of the first order autoregressive parameter, *Journal of the American Statistical Association*, **74**, 140–146.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987–1007.

Franses, P. H. and Ghijsels, H. (1999). Additive outliers, GARCH and forecasting volatility, *International Journal of Forecasting*, **15**, 1–9.

Godambe, V. P. (1985). The foundations of finite sample estimation in stochastic processes, *Biometrika*, **72**, 419–428.

Huber, P. J. (1981). *Robust Statistics*, John wiley & Sons, New York.

Kim, S. and Hwang, S. Y. (2007). A doubly robustified estimating function for ARCH time series models, *Journal of the Korean Statistical Society*, **36**, 387–396.

The Robust Estimation Method for Analyzing the Financial Time Series Data

S. Kim[1]

[1]Dept. of Statistics, Chung-Ang University

(Received April 2008; accepted May 2008)

Abstract

In this paper, we propose the double robust estimators which are the solutions of the double robust estimating equations to analyze and treat the outliers in the stock market data in Korea including the IMF period. The feasibility study shows that the proposed estimators work quitely better than the least squares estimators and the conventional robust estimators.

Keywords: Robust estimators, GARCH model, stock indexes.

This research was supported by the Chung-Ang University Grants in 2007.

[1]Associate Professor, Dept. of Statistics, Chung-Ang University, Dongjak-Gu, Seoul 156-756, Korea.
E-mail: sahm@cau.ac.kr