

변환된 GARCH모형에서의 예측값 추정

박주연¹ · 여인권²

¹숙명여자대학교 통계학과, ²숙명여자대학교 통계학과

(2009년 7월 접수, 2009년 7월 채택)

요약

이 논문에서는 GARCH 모형에서 변환-역변환 방법을 통해 예측값을 추정할 때 발생하는 편향을 줄이기 위한 방법을 소개한다. 모수적 붓스트랩을 활용하여 본래 척도에서의 최소평균제곱오차 예측값인 조건부 기대값을 계산한다. KOSPI와 KOSDAQ 수익률 분석을 통해 제안한 방법이 편향을 줄여주는 것을 확인하였다.

주요용어: 모수적 붓스트랩, 최소평균제곱오차 예측, Yeo-Johnson 변환.

1. 서론

일반적으로 시계열 자료 분석은 시점에 관계없이 시계열의 평균과 분산이 일정하고 공분산은 시차에만 영향을 받는다는 정상성 가정 하에서 이루어진다. 자료가 정상성을 만족하지 않는 경우, 차분이나 변환 등을 통해 자료가 정상성을 만족하도록 만든 후 정상시계열 모형에 적용하고 있다. 1980년대 초까지의 시계열분석에서는 조건부 1차 적률의 시간종속성을 모형화 하는 것이 주된 관심사였고 분산은 장애모수(nuisance parameter)로 취급되고 있었다. 주식, 선물, 옵션, 채권, 예금 등의 투자 상품에 대한 수익률이나 각종 국가경제 지표 등의 자료들은 대부분 사회경제적 요인의 변화에 따른 영향에 민감하게 반응하기 때문에 앞 시점들의 관측값에 따라 변동의 폭이 다르게 나타나는 경향이 있다. 자료가 정상성을 만족하여 비조건부 분산은 시간에 대해 불변이지만 과거자료의 정보가 주어진 조건부 분산이 과거 자료에 영향을 받는 자료를 통계학에서는 조건부이분산시계열자료(conditional heteroscedastic time series)라고 부른다.

이분산시계열에 대한 연구는 Engle (1982)의 autoregressive conditional heteroscedastic(ARCH) 모형과 Nichollas와 Quinn (1982)의 random coefficient autoregressive모형이 제안되면서 본격적으로 연구가 시작되었다. 이후 ARCH모형을 개선하고 확장시킨 모형들이 소개되었는데 Bollerslev (1986)는 고차의 차수를 가지는 ARCH모형을 보다 단순하게 표시할 수 있는 generalized ARCH(GARCH) 모형을, Nelson (1990)은 변동에 갑작스러운 변화가 있는 현상을 설명하기 위한 integrated-GARCH 모형을, Engle 등 (1987)는 시계열자료의 평균이 그 시점에서의 변동에 영향을 받는 현상을 설명하기 위한 GARCH-M 모형을, Higgins와 Bera (1992)은 Box-Cox변환을 이용한 power-transformed ARCH 등을 제안하였다. 이러한 시계열 분석 모형은 주로 정규성 또는 대칭성(symmetry)을 가정하고 있으나, Fama (1965), Rabemanjara와 Zakoian (1993), McCulloch (1996), McDonald (1996) 등이 언급

본 연구는 숙명여자대학교 2008년도 교내연구비 지원에 의해 수행되었음.

²교신저자: (140-742) 서울특별시 용산구 청파동 2가, 숙명여자대학교 통계학과, 부교수.

Email: inkwon@sm.ac.kr

한 것과 같이 금융 시계열들은 비대칭인 형태를 가지고 있으며, 특히 조건부 이분산의 형태가 좋은 소식(good news) 보다 나쁜 소식(bad news)에 훨씬 더 영향을 받는 레버리지 효과(leverage effect)를 가지는 것으로 알려졌다. 비대칭 효과는 변동(volatility)의 절대값이 동일하다고 하더라도 양의 변동보다 음의 변동이 더 큰 이분산 효과를 가지는 것으로, 감소추세가 있는 시계열에서의 위험 또는 변동성이 더 크게 느껴진다는 것을 의미한다.

조건부 이분산에서의 비대칭 현상을 해결하기 위한 여러 방안이 제안되었는데 조건부 분산식을 비대칭적으로 모형화하는 것이 주요 방법 중 하나이다. 대표적인 비대칭 조건부 이분산 모형으로는 Nelson (1991)의 exponential-GARCH모형, Glosten 등 (1993)의 GJR-GARCH 모형, Ding 등 (1993)의 asymmetric power ARCH모형, Hwang과 Kim (2004)의 power-transformed threshold ARCH모형 등이 있다. 두 번째 방법은 시계열자료의 오차항의 분포를 정규분포나 t -분포와 같은 대칭인 형태를 가지는 분포 대신 비대칭적인 형태를 가지는 분포로 가정하는 것으로 Fernandez와 Steel (1998)의 비대칭 t -분포(skewed student- t distribution)가 대표적인 예이다. 또 다른 방법으로는 자료변환을 통해 시계열의 정규성 또는 대칭성을 높인 후 모형을 추정하는 방법이다. 비대칭 모형이나 비대칭 분포를 가정하는 경우, 모수를 추정하는데 어려움이 있을 뿐만 아니라 해석이나 응용이 쉽지 않는 단점이 있다.

시계열 분석에 있어 궁극적인 목적은 현재까지 관측된 시계열의 특성을 파악하고 이를 이용하여 미래에 얻어질 자료에 대해 예측하는 것이다. 일반적으로 많이 사용되고 있는 예측값은 평균제곱오차를 최소로 만드는 값으로 추정하는데 이 값은 과거 시계열 자료의 조건부 기대값으로 표시된다. 변환을 통해 얻어진 예측은 변환된 척도로 표시되기 때문에 본래 시계열의 척도에서의 예측값을 얻기 위해서는 역변환을 해야 한다. 문제는 비선형 변환-역변환을 통해 예측값을 추정할 경우, 기대값을 변환한 값과 변환된 확률변수의 기대값이 서로 다르기 때문에 변환-역변환을 이용한 예측에서 편향(bias)이 발생하게 된다는 것이다. 이 논문에서는 GARCH모형에서 변환-역변환을 이용하여 예측값을 구했을 때 편향이 발생하는 문제를 해결하기 위한 방법을 제시하고 KOSPI와 KOSDAQ의 수익률 분석을 통해 제안한 방법의 타당성을 확인하고자 한다.

2. 변환된 GARCH모형

2.1. 변환선택

금융시계열 자료는 서론에서 언급했던 것과 같이 대부분 비대칭적인 특성을 가지기 때문에 분석이 복잡하고 통계적 추론이 어렵다. 비대칭성으로 인해 정규성을 만족하지 않는 경우, Box-Cox 변환 (Box와 Cox, 1964)을 이용하여 자료의 정규성을 높인 후 분석하는 방법들이 많이 연구되었다. 하지만 Box-Cox 변환은 자료의 값이 양수인 경우에만 사용할 수 있기 때문에 수익률과 같이 0 또는 음수가 발생하는 금융시계열 자료분석에서는 적용이 어렵다. 이 논문에서는 모든 실수값에 적용시킬 수 있는 Yeo-Johnson 변환 (Yeo와 Johnson, 2000)을 사용하기로 한다. Yeo-Johnson 변환은 다음과 같이 정의되는데

$$\psi(x, \lambda) = \begin{cases} \frac{(x+1)^\lambda - 1}{\lambda}, & \lambda \neq 0, x \geq 0, \\ \log(x+1), & \lambda = 0, x \geq 0, \\ -\frac{(-x+1)^{2-\lambda} - 1}{(2-\lambda)}, & \lambda \neq 2, x < 0, \\ -\log(-x+1), & \lambda = 2, x < 0, \end{cases}$$

$\lambda > 1$ 이면 전체적으로 볼록하기 때문에 음의 왜도 문제를 해결할 수 있고 $\lambda < 1$ 이면 오목한 형태를 보이고 있으므로 양의 왜도를 줄일 수 있다. 일반적으로 음의 변동에 더 영향을 많이 받는 금융 시계열 분

석에서는 대부분 음의 왜도를 가지며 따라서 $\lambda > 1$ 인 경우가 많다. 변환-역변환 방법에 적용하기 위한 Yeo-Johnson 변환의 역변환 함수는 다음과 같다.

$$\psi^{-1}(z, \lambda) = \begin{cases} (\lambda z + 1)^{\frac{1}{\lambda}} - 1, & z \geq 0, \lambda \neq 0, \\ \exp(z) - 1, & x \geq 0, \lambda < 0, \\ 1 - \{(\lambda - 2)z + 1\}^{\frac{1}{2-\lambda}}, & x < 0, \lambda \neq 2, \\ 1 - \exp(-z), & x < 0, \lambda = 2. \end{cases}$$

2.2. 모수 추정

Engle (1982)에 의해 ARCH(autoregressive conditional heteroscedastic) 모형이 소개되면서 조건부이 분산시계열자료에 대한 연구가 본격적으로 시작되었다. 시점 t 에서의 오차항을 ε_t 이라고 하고 I_{t-1} 를 $t-1$ 시점까지의 시계열 정보라고 하면, 조건부 분산은 $h_t = \text{Var}(\varepsilon_t | I_{t-1})$ 로 정의된다. ARCH(p) 모형에서는 조건부분산의 구조식이 다음과 같이 p -시차 이전의 제곱오차의 선형결합으로 표시된다고 가정한다.

$$h_t = \alpha_0 + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2,$$

여기서 $\alpha_0 > 0$ 이고 $\alpha_j \geq 0$ 이며 α_j 는 ε_t 의 비조건부 분산을 유한하게 만드는 조건을 만족해야 한다. Bollerslev (1986)는 고차의 차수를 가지는 ARCH모형을 보다 단순하게 표시할 수 있도록 ARCH 모형에서 다음과 같이 이전 조건부 분산 항을 추가한 GARCH(Generalized ARCH) 모형을 제안하였다.

$$h_t = \alpha_0 + \sum_{j=1}^p \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^q \beta_j h_{t-j},$$

여기서 $\alpha_0 > 0$, $\alpha_j \geq 0$, $\beta_j \geq 0$ 이며 $\sum_{j=1}^{\max(p,q)} (\alpha_j + \beta_j) < 1$ 를 만족한다. 앞에서도 언급한 것과 같이 ARCH모형이나 GARCH모형은 조건부분산식에서 ε^2 을 사용하기 때문에 절대값이 같은 음과 양의 오차에 대해 동일한 영향력을 받도록 되어 있어 레버리지 효과를 설명할 수 없다. 이 논문에서는 이를 해결하기 위해 시계열 $\{X_t\}_{t=1}^n$ 를 변환시킨 시계열 $\{\psi(X_t, \lambda)\}_{t=1}^n$ 이 다음과 같은 ARMA(a, b)-GARCH(p, q) 모형을 따른다고 가정한다.

$$\Phi(B) \{\psi(X_t, \lambda) - \mu\} = \Theta(B)\varepsilon_t, \quad \varepsilon_t = \sqrt{h_t}\varepsilon_t, \tag{2.1}$$

여기서 ε_t 는 분산이 1인 백색잡음을 나타내고 $\Phi(B)$ 와 $\Theta(B)$ 는 각각 AR작용소와 MA작용소를 나타내는데 후진작용소 B 로 표시하면 다음과 같다.

$$\Phi(B) = 1 - (\phi_1 B + \phi_2 B^2 + \dots + \phi_a B^a), \quad \Theta(B) = 1 - (\theta_1 B + \theta_2 B^2 + \dots + \theta_b B^b)$$

추정할 모수를 $\xi = (\mu, \phi, \theta, \alpha, \beta, \lambda)$ 라고 하면, ε_t 가 정규분포를 따른다는 가정 하에서, 로그가능도함수(log likelihood function)는 다음과 같다.

$$L(\xi | x) = \sum_{t=1}^n \left[-\frac{1}{2} \log(h_t) - \frac{1}{2h_t} [\Theta(B)^{-1} \Phi(B) \{\psi(x_t, \lambda) - \mu\}]^2 + \log(J(x_t, \lambda)) \right],$$

여기서 $J(x_t, \lambda)$ 는 변환된 확률변수들에 대한 자코비안으로 Yeo-Johnson 변환인 경우 다음과 같이 계산된다.

$$J(x, \lambda) = \frac{\partial}{\partial x} \psi(x, \lambda) = (|x| + 1)^{\text{sign}(x)(\lambda-1)}.$$

위의 로그가능도 함수를 최대로 만드는 최대가능도 추정값 $\hat{\xi} = (\hat{\mu}, \hat{\phi}, \hat{\theta}, \hat{\alpha}, \hat{\beta}, \hat{\lambda})$ 을 수치해석학적인 방법을 통해 계산할 수 있다. 모수 λ 가 1에서 멀리 떨어져 있으면 있을수록 레버리지 효과가 더욱더 강하다고 볼 수 있기 때문에 추정된 $\hat{\lambda}$ 를 통해 레버리지 효과에 대한 검정을 쉽게 수행할 수 있다.

3. 예측값 추정

시계열 분석의 주요 목적 중 하나는 지금까지 관측된 시계열의 패턴이 미래에도 지속된다는 가정에서 미래에 대한 예측을 하는 것이다. 이러한 예측에서 가장 많이 사용되는 최소평균제곱오차(Minimum Mean Square Error, MMSE) 예측은 시점 t 에서의 시계열 자료 Z_t 에 대한 예측값으로 평균제곱오차 $E[Z_t - \hat{Z}_t]^2$ 를 최소로 만드는 \hat{Z}_t 를 사용한다. 변환된 시계열을 $Z_t = \psi(X_t, \lambda)$ 라고 표시하고 시점 n 에서 l 시차 후의 시계열 Z_{n+l} 에 대한 MMSE 예측값을 $Z_n(l)$ 이라고 하면, $Z_n(l)$ 은 조건부 기대값 $E[Z_{n+l}|Z_n, Z_{n-1}, \dots]$ 가 되는데 이 기대값을 계산하는 과정에서, $j > 0$ 에 대해, $E[\varepsilon_{n+j}|Z_n, Z_{n-1}, \dots] = 0$ 이 되기 때문에 $Z_n(l)$ 은 다음과 같이 비교적 간단한 형태로 유도된다.

$$Z_n(l) = \begin{cases} \mu + \sum_{j=1}^a \phi_j \{Z_n(l-j) - \mu\} - \sum_{j=1}^b \theta_j \varepsilon_{n+l-j}, & l \leq b, \\ \mu + \sum_{j=1}^a \phi_j \{Z_n(l-j) - \mu\}, & l > b. \end{cases} \quad (3.1)$$

이를 바탕으로 $Z_n(l)$ 의 추정량은 앞에서 구한 최대가능도 추정량 $\hat{\xi}$ 를 이용하여 다음과 같이 구한다.

$$\hat{Z}_n(l) = \begin{cases} \hat{\mu} + \sum_{j=1}^a \hat{\phi}_j \{\hat{Z}_n(l-j) - \hat{\mu}\} - \sum_{j=1}^b \hat{\theta}_j \hat{\varepsilon}_{n+l-j}, & l \leq b, \\ \hat{\mu} + \sum_{j=1}^a \hat{\phi}_j \{\hat{Z}_n(l-j) - \hat{\mu}\}, & l > b, \end{cases}$$

여기서 $j \geq 0$ 에 대해, $\varepsilon_{n-j} = Z_{n-j} - \hat{Z}_{n-j}$ 는 $n-j$ 시점에서의 잔차이고 $\hat{Z}_n(l-j)$ 는 Z_n, Z_{n-1}, \dots 와 $\hat{\varepsilon}_n, \hat{\varepsilon}_{n-1}, \dots$ 로 표시될 때까지 위의 관계식을 재귀적으로 계산하여 구할 수 있다.

변환된 척도가 아닌 본래 척도에서의 $n+l$ 시점의 예측값은 변환된 척도에서의 예측값에 역변환을 이용한 $\psi^{-1}(\hat{Z}_n(l), \hat{\lambda})$ 으로 추정하는 것이 일반적인 방법이다. 문제는 이러한 방법으로 예측값에 대한 추정량을 구했을 때 그 추정량은 MMSE 예측값의 추정량이 아닌 편향된 예측값의 추정량이라는 것이다. 본래 척도에서 n 시점에서 l 시차 후의 시계열 X_{n+l} 에 대한 MMSE 예측값은 $E[X_{n+l}|X_n, X_{n-1}, \dots]$ 이고 이 예측값을 변환된 척도에서의 시계열로 표시하면 다음과 같다.

$$\begin{aligned} X_n(l) &= E[X_{n+l}|X_n, X_{n-1}, \dots] = E[\psi^{-1}(Z_{n+l}, \lambda)|Z_n, Z_{n-1}, \dots] \\ &= E \left[\psi^{-1} \left(\mu + \sum_{j=1}^a \phi_j \{Z_n(l-j) - \mu\} - \sum_{j=1}^b \theta_j \varepsilon_{n+l-j}, \lambda \right) \middle| Z_n, Z_{n-1}, \dots \right], \end{aligned} \quad (3.2)$$

여기서 λ 가 1이 아니면 ψ^{-1} 는 비선형함수가 되기 때문에, ψ^{-1} 안에 있는 n 시점 이후의 예측오차항들이 예측식에 여전히 남아 있어 기대값 계산이 어렵게 된다.

변환-역변환 방법을 이용한 예측값에 있어 발생하는 편향 문제를 해결하기 위해, Taylor 전개를 통해 편향을 추정하여 예측값에서 이 편향을 빼주는 방법과 기대값을 직접 추정하는 방법을 고려할 수 있다. 편향을 예측할 경우 ψ^{-1} 의 도함수 계산과 분산 추정을 해야 하는 번거로움이 있기 때문에 이 논

문에서는 기대값을 직접 추정하는 방법에 대해 알아본다. 유사연구로는 회귀분석에서 Duan (1983)의 smearing 추정과 ARMA모형에서 여인권과 조혜민 (2008) 등이 있다. 기대값 (3.2)에서, $j > 1$ 에 대해, $Z_n(l-j)$ 는 $Z_{n+l-j-1}, Z_{n+l-j-2}, \dots$ 와 $\varepsilon_{n+l-j}, \varepsilon_{n+l-j-1}, \dots$ 으로 표시되고 이를 재귀적으로 계산하여 정리하면 $X_n(l)$ 는 Z_n, Z_{n-1}, \dots 와 $\varepsilon_{n+l}, \varepsilon_{n+l-1}, \dots$ 로 이루어진 식의 기대값으로 표시될 수 있다. 여기서, $t \leq n$ 에 대해, 조건부 기대값 안의 ε_t 는 $\hat{\varepsilon}_t$ 로 대체할 수 있으며 조건부 기대값 (3.2)을 추정하기 위해서는 $\varepsilon_{n+l}, \dots, \varepsilon_{n+1}$ 으로 이루어진 식의 기대값을 추정해야 한다. 예측오차 ε_t 는 위의 가정 (2.1)에 의하면 $N(0, h_t)$ 를 따르기 때문에 h_t 를 추정하고 이 분포에서 난수를 반복적으로 B 번 발생시켜 다음과 같은 표본평균으로 조건부 기대값을 추정할 수 있다.

$$\hat{X}_n(l) = \frac{1}{B} \sum_{b=1}^B \psi^{-1} \left(\hat{\mu} + \sum_{j=1}^a \hat{\phi}_j (Z_{n+l-j}^* - \hat{\mu}) - \sum_{j=1}^b \hat{\theta}_j \varepsilon_{n+l-j}^*, \hat{\lambda} \right), \quad (3.3)$$

여기서 ε_t^* 는 $t \leq n$ 일 때 잔차 $\hat{\varepsilon}_t$ 로 대신하고 $t > n$ 일 때는 $N(0, \hat{h}_t)$ 에서 발생시킨 난수로 대신하는데 \hat{h}_t 는 다음과 같이 계산된다.

$$\hat{h}_t = \hat{\alpha}_0 + \sum_{j=1}^p \hat{\alpha}_j \varepsilon_{t-j}^2 + \sum_{j=1}^q \hat{\beta}_j \hat{h}_{t-j}, \quad \varepsilon_{t-j}^2 = \begin{cases} \hat{\varepsilon}_{t-j}^2, & t-j \leq n, \\ \hat{h}_{t-j}, & t-j > n \end{cases}$$

그리고 시계열 Z_t^* 는 Z_n, Z_{n-1}, \dots 와 $\varepsilon_t^*, \varepsilon_{t-1}^*, \dots$ 로 표현되도록 재귀적으로 전개하여 구할 수 있다.

4. 실증분석

이 절에서는 KOSPI지수와 KOSDAQ지수의 로그수익률에 대해 조건부이분산성과 레버리지 효과가 있는지를 알아보고 원래 척도에서 예측한 경우(방법 I), 변환-역변환을 통해 예측한 경우(방법 II), 앞에서 언급한 방법을 통해 예측한 경우(방법 III)를 비교한다. P_t 를 t -시점에서의 주가지수라고 하면 로그수익률은 다음과 같이 정의한다.

$$Z_t = 100 \times \log \left(\frac{P_t}{P_{t-1}} \right).$$

실증분석에서는 아래와 같이 $n = 100, 200, 300$ 개의 시계열 자료로 이루어진 블록에 대해 모형을 추정하고 이 추정된 모형을 근거로 5시차 후까지의 예측값을 추정한다. 이동평균을 계산할 때처럼 시계열을 한 칸씩 이동시키면서 같은 작업을 각각 500번을 반복한다.

$$Z_t = \{Z_{t+1}, \dots, Z_{t+n}\}, \quad t = 0, \dots, 499,$$

여기서 첫 번째 자료 Z_1 는 2005년 1월 4일의 수익률을 나타낸다. Matlab 2007을 사용하여 분석하였는데 원래 척도에서의 예측은 garchfit 함수를 이용하여 상수항-GARCH(1,1), AR(1)-GARCH(1,1), MA(1)-GARCH(1,1)을 적합시킨 후 AIC로 세 모형 중 최적의 모형을 선택하고 garchpred 함수를 이용하여 예측값을 추정하였다. 변환-역변환의 경우, 세 모형 각각에 대해 garchfit 함수의 로그가능도함수 값에 Jacobian의 로그를 추가하여 fminbnd 함수으로 변환모수 $\hat{\lambda}$ 을 구하고 이에 따른 모형의 모수를 추정하였다. 원래 척도에서의 예측과 마찬가지로 AIC로 최적의 모형을 선택한 후 변환된 척도에서의 예측값을 garchpred 함수로 추정하고 역변환을 적용하여 예측값을 구했다. 제안된 방법은 변환된 척도에서의 예측값에 표준정규분포에서 5개의 난수를 B 번 발생시키고 garchpred 함수에서 계산된 $\sqrt{\hat{h}_t}$ 를 곱하여 얻은 $\varepsilon_{t+n+1}^*, \dots, \varepsilon_{t+n+5}^*$ 를 식 (3.3)에 대입하여 예측값을 추정하였다.

그림 4.1은 분석에 사용된 800개의 수익률 전체에 대해 ARCH 효과가 있는지를 알아보기 위해 수익률의 ACF와 수익률의 제곱에 대한 자기상관함수(ACF)와 부분자기상관함수(PACF)를 표시한 것이다.

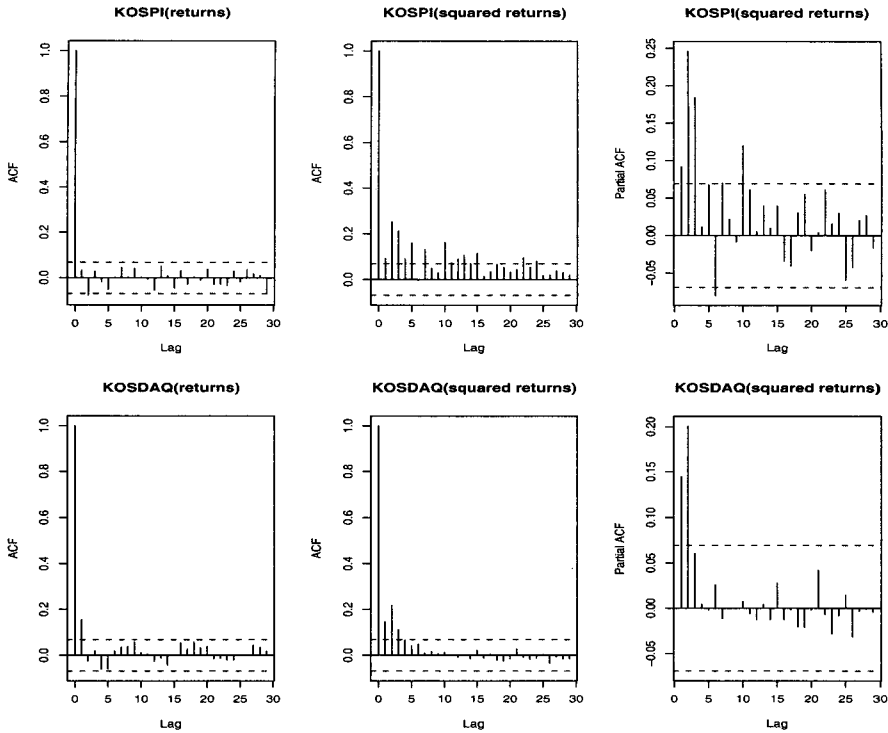


그림 4.1. KOSPI와 KOSDAQ 수익률에 대한 자기상관분석

KOSPI의 경우, 수익률에 대해서는 자기상관이 없으나 제곱수익률에 대해서는 자기상관을 가지는 전형적인 ARCH 효과를 보이고 있으며 KOSDAQ의 수익률은 1차 자기상관을 가지는 것으로 나타났으나 MA(1)모형 적합 이후에 대한 잔차의 ARCH 효과가 있는 것으로 나타났다. 차수를 12까지 했을 때 라그랑지승수 검정(Lagrange multiplier test)에서도 통계량(유의확률)이 각각 104.136(<0.001)과 51.587(<0.001)로 나와 ARCH효과가 있는 것을 확인하였다. 레버리지 효과 여부에 대한 검정은 각각의 n 에 대해 얻어지는 500개의 $\hat{\lambda}$ 를 이용하여 분석하였다. 각 블록에 대한 λ 의 추정값의 패턴은 그림 4.2와 같이 주어졌는데 두 지수의 수익률에 대한 $\hat{\lambda}$ 들이 대부분 1보다 커 레버리지 효과가 있는 것으로 나타났다. 주어진 기간 동안 KOSPI의 경우 $n = 100, 200, 300$ 일 때 $\hat{\lambda}$ 의 평균은 각각 1.184, 1.184, 1.186인 것으로 나타났으며, KOSDAQ의 경우 1.348, 1.345, 1.323인 것으로 나타나 KOSDAQ의 수익률이 KOSPI의 수익률보다 레버리지 효과가 더 큰 것으로 분석되었다.

표 4.1은 1-시차 예측오차와 1-시차에서 5-시차까지의 예측오차 전체에 대한 평균과 표준편차를 정리한 것이다. 이 표에 의하면 모든 n 에서 방법 III의 예측오차 평균의 절대값, 즉 편향이 가장 작은 것으로 나타났으며 표준편차는 KOSPI 수익률의 경우 방법 I이, KOSDAQ 수익률의 경우 방법 II가 작은 것으로 나타났으나 편향에 비해 방법간의 차이는 크지 않은 것으로 나타났다. 흥미로운 것은 n 이 200일 때 KOSDAQ 수익률에 대한 방법 III을 제외하고는 n 이 커질수록 편향과 표준편차가 커지는 경향이 있다는 것이다. 이것은 시계열의 특성이 일정기간 동안 비슷하게 유지되지만 장기적으로 그 특성이 지속되지 않는다는 것을 의미하며 수익률을 분석하는데 있어 이러한 특성을 고려한 모형이나 적절한 시계열 크기의 선택에 대한 연구가 요구된다.

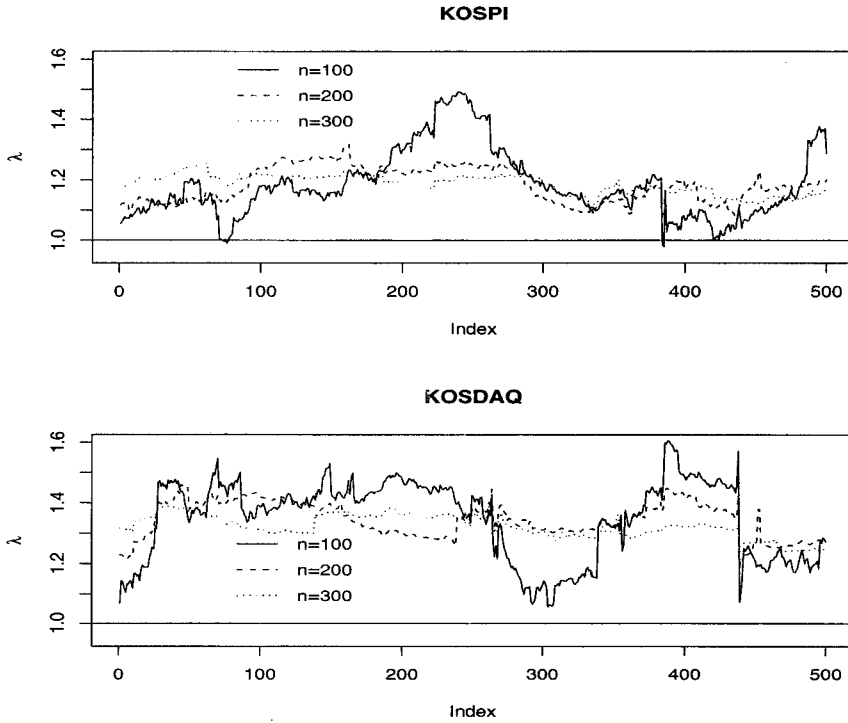


그림 4.2. KOSPI와 KOSDAQ 수익률에 대한 λ 의 패턴

표 4.1. 예측오차의 평균과 표준편차

n	방법	KOSPI				KOSDAQ			
		1-시차		전체		1-시차		전체	
		평균	표준편차	평균	표준편차	평균	표준편차	평균	표준편차
100	I	0.006	1.086	0.005	1.088	0.046	1.360	0.058	1.358
	II	0.079	1.093	0.077	1.092	0.151	1.338	0.159	1.358
	III	0.004	1.095	0.000	1.092	-0.030	1.386	-0.028	1.367
200	I	0.030	1.257	0.029	1.256	0.081	1.496	0.103	1.500
	II	0.103	1.259	0.102	1.258	0.194	1.484	0.209	1.500
	III	0.021	1.260	0.017	1.261	0.007	1.516	0.005	1.519
300	I	0.090	1.382	0.086	1.382	0.142	1.434	0.159	1.429
	II	0.163	1.382	0.159	1.382	0.244	1.424	0.257	1.427
	III	0.070	1.387	0.062	1.386	0.067	1.459	0.068	1.451

5. 결론

일반적인 시계열분석에서는 시계열 오차항이 정규분포 또는 t -분포와 같이 대칭적인 분포를 따른다는 가정 하에서 추론이 이루어진다. 시계열이 이러한 가정을 만족하지 않는 경우 쉽게 적용할 수 있는 방법이 시계열자료를 가정에 근사하도록 변환하는 것이다. 추정된 모형을 근거로 본래척도에서의 예측값을 구하려고 할 때 흔히 변환된 척도에서의 예측값을 역변환하여 본래척도의 예측값을 얻는다. 문제는

예측값 추정을 위해 사용되는 MMSE 예측값은 조건부 기대값으로 표시되며, 비선형 변환에 대해 변환된 변수의 기대값이 변수의 기대값을 변환한 값과 일치하지 않는 것처럼, 변환된 자료로부터 얻은 예측 추정값을 역변환하여 본래 척도에서의 예측추정값을 계산하면 이 값은 MMSE 예측값이 되지 못하고 항상 편향이 발생한다는 것이다. 이 논문에서는 변환-역변환을 이용하여 GARCH 모형에서의 예측할 때 발생하는 편향의 문제를 모수적 붓스트랩을 통해 해결하는 방안에 대해 알아보았다. 또한 KOSPI와 KOSDAQ의 수익률에 대한 실증분석을 통해 제안된 방법이 편향을 줄여주는 것을 확인하였다.

참고문헌

- 여인권, 조혜민 (2008). 변환-역변환을 통한 자기회기이동평균모형에서의 예측값 추정, <응용통계연구>, **21**, 537-546.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307-328.
- Box, G. E. P. and Cox, C. R. (1964). An analysis of transformations, *Journal of the Royal Statistical Society, Series B*, **26**, 211-252.
- Ding, Z., Granger, C. W. J. and Engle, R. F. (1993). A long memory property of stock market returns and a new model, *Journal of Empirical Finance*, **1**, 83-106.
- Duan, N. (1983). Smearing estimate: A nonparametric retransformation method, *Journal of the American Statistical Association*, **78**, 605-610.
- Fama, E. F. (1965). The behavior of stock market prices, *The Journal of Business*, **38**, 34-105.
- Fernandez, C. and Steel, M. F. J. (1998). On Bayesian modelling of fat tails and skewness, *Journal of the American Statistical Association*, **93**, 359-371.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987-1008.
- Engle, R. F., Lilien, D. and Robins, R. P. (1987). Estimating time varying risk premia in the term structure: The ARCH-M model, *Econometrica*, **55**, 391-407.
- Glosten, L. R., Jagannathan, R. and Runkle, D. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks, *Journal of Finance*, **48**, 1779-1801.
- Higgins, M. L. and Bera, A. K. (1992). A class of nonlinear ARCH models, *International Economic Review*, **33**, 137-158.
- Hwang, S. Y. and Kim, T. Y. (2004). Power transformation and threshold modeling for ARCH innovations with applications to tests for ARCH structure, *Stochastic Processes and Their Applications*, **110**, 295-314.
- McCulloch, J. H. (1996). Financial applications of stable distributions, in statistical methods in finance, *Handbook of Statistics*, **14**, 393-425, North-Holland, Amsterdam.
- McDonald, J. B. (1996). Probability distribution for financial models, *Handbook of Statistics*, **14**, 427-461, North-Holland, Amsterdam.
- Nelson, D. B. (1990). Stationarity and persistence in the GARCH(1,1) model, *Econometric Theory*, **6**, 318-334.
- Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach, *Econometrica*, **59**, 347-370.
- Nichollas, D. F. and Quinn, B. G. (1982). Random coefficient autoregressive models: An introduction, *Lecture Notes in Statistics*, **11**, Springer-Verlag, New York.
- Rabemanjara, R. and Zakoian, J. M. (1993). Threshold ARCH model and asymmetries in volatility, *Journal of Applied Econometrics*, **8**, 31-49.
- Yeo, I. K. and Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry, *Biometrika*, **87**, 954-959.

Prediction Value Estimation in Transformed GARCH Models

Juyeon Park¹ · In-Kwon Yeo²

¹Department of Statistics, Sookmyung Women's University

²Department of Statistics, Sookmyung Women's University

(Received July 2009; accepted July 2009)

Abstract

In this paper, we introduce the method that reduces the bias when the transformation and back-transformation approach is applied in GARCH models. A parametric bootstrap is employed to compute the conditional expectation which is the prediction value to minimize mean square errors in the original scale. Through the analysis of returns of KOSPI and KOSDAQ, we verified that the proposed method provides a bias-reduced estimation for the prediction value.

Keywords: Minimum mean square error, parametric bootstrap, Yeo-Johnson transformation.

This research was supported by the Sookmyung Women's University Research Grants 2008.

²Corresponding author: Associate Professor, Department of Statistics, Sookmyung Women's University, Seoul 140-742, Korea. E-mail: inkwon@sm.ac.kr