

Multivariate volatility for high-frequency financial series

G.J. Lee^a · Sun Young Hwang^{a,1}

^aDepartment of Statistics, Sookmyung Women's University

(Received December 23, 2016; Revised January 18, 2017; Accepted January 18, 2017)

Abstract

Multivariate GARCH models are interested in conditional variances (volatilities) as well as conditional correlations between return time series. This paper is concerned with high-frequency multivariate financial time series from which realized volatilities and realized conditional correlations of intra-day returns are calculated. Existing multivariate GARCH models are reviewed comparatively with the realized volatility via canonical correlations and value at risk (VaR). Korean stock prices are analysed for illustration.

Keywords: high-frequency time series, realized volatility, multivariate GARCH

1. 서론

금융시계열에서 변동성(volatility)은 ‘일정기간 주식, 채권, 또는 상품의 가격이 변동하는 정도’를 의미한다. 금융시장에서 수익률의 변동성은 일반적인 시계열과 달리 시간에 따라 변하는 특성을 가지며 특정기간동안 큰 값이나 작은 값이 지속되는 변동성 집중(volatility cluster) 현상, 가격의 급락과 급등에 다르게 반응하는 비대칭 효과(leverage effect) 등의 특징을 가진다. 변동성은 불확실성(uncertainty)을 나타내기 때문에 위험(risk)의 측도가 되며 옵션의 가격결정, 포트폴리오의 선택 등에 있어 중요한 역할을 한다 (cf. Kim과 Lee, 2005; Yoon과 Hwang, 2015). Engle (1982)은 변동성을 과거의 관측 값들로 모형화 하는 자기회귀 조건부 이분산(ARCH) 모형을 제안하였고, Bollerslev (1986)는 일반화된 자기회귀 조건부 이분산(GARCH) 모형으로 ARCH 모형을 일반화하였다.

수익률의 변동성은 다른 수익률의 변동에 영향을 미치는 것으로 알려져 있다. 따라서 변동성간에는 동적인 관계(dynamic correlation)가 존재하며 이러한 특성을 반영하여 모델링이 수행되어야한다. 다변량 수익률 변동성의 동적인 관계를 구조화한 모형에는 단변량 GARCH 모형의 단순 확장 형태인 exponential weighted moving average(EWMA) 모형, diagonal VEC(DVEC) 모형 및 Baba-Engle-Kraft-Kroner(BEKK) 모형 등이 있다. 단순 확장 모형은 차원이 증가함에 따라 모수가 급격하게 많아진다는 단점이 있으므로 이를 해결하기 위한 모형으로 상관계수가 시점에 관계없이 일정하다고 가정하여 모수의 개수를 줄인 constant conditional correlation(CCC) 모형과 CCC 모형의 일반화된 형태인 dynamic conditional correlation(DCC) 모형 등이 있다 (Choi 등, 2009; Hwang 등, 2009).

This research was supported by Basic Science Research Program through the NRF of Korea funded by the Ministry of Education (2015-057031).

¹Corresponding author: Department of Statistics, Sookmyung Women's University, 100, Cheongpa-ro 47-gil, Yongsan-gu, Seoul 04310, Korea. E-mail: shwang@sookmyung.ac.kr

최근에는 데이터 처리능력의 향상으로 고빈도 자료(high-frequency data)를 이용하여 변동성을 추정하는 방법이 고안되고 있다. Andersen과 Bollerslev (1997)와 Anderson 등 (2003)은 일중수익률의 변동성을 추정하는 방법으로 실현변동성을 제안하였다. 실현변동성(realized volatility; RV)이란 고빈도 자료로부터의 일중 수익률(intra-day return)로 계산한 일중변동성을 의미한다. 실현변동성은 과거 변동성의 예측을 평가하는 중요한 역할을 한다.

본 연구에서는 다변량 GARCH 모형과 실현변동성을 소개하고 국내주가지료를 통해 실증분석해보고자 한다. 1분, 5분 단위의 고빈도 자료로 계산한 실현변동성을 참 값으로 간주하고 정준상관(canonical correlation) 분석과 VaR-분석을 통해 다변량 GARCH 모형들을 비교평가해보고자 한다.

2. Multivariate GARCH(MGARCH) 모형

본 절에서는 다변량 변동성을 모형화하는 MGARCH 모형에 대해 소개한다. 기본적인 개념과 수식은 Hwang 등 (2009)과 Tsay (2010)를 참고하여 정리하였다. k 개의 수익률로 구성된 벡터 $\mathbf{r}_t = (r_{1t}, \dots, r_{kt})^T$ 를 다음과 같이 모형화 한다.

$$\mathbf{r}_t = \boldsymbol{\mu}_t + \mathbf{a}_t.$$

$\boldsymbol{\mu}_t$ 는 평균수익률 벡터이며, 일반적으로 VARMA(vector ARMA) 모형으로 적합 시킨다. 오차항 $\mathbf{a}_t = (a_{1t}, \dots, a_{kt})^T$ 는 다변량-GARCH 모형으로서 다음과 같다.

$$\mathbf{a}_t = \mathbf{H}_t^{\frac{1}{2}} \mathbf{e}_t,$$

여기서 \mathbf{H}_t 는 $k \times k$ 양정치 행렬이고 \mathbf{e}_t 는 $k \times 1$ 벡터로 다음의 두 조건을 만족한다.

$$(i) E(\mathbf{e}_t) = \mathbf{0}, \quad (ii) \text{Var}(\mathbf{e}_t) = \mathbf{I}_k.$$

위 조건을 만족하면 $t - 1$ 시점까지의 정보가 주어졌을 때 \mathbf{a}_t 의 조건부 분산-공분산행렬은 $\mathbf{H}_t = \text{Cov}(\mathbf{a}_t | F_{t-1})$ 이 된다. 이때, F_{t-1} 은 $t - 1$ 시점까지 포함된 정보 집합을 나타낸다. 다변량 변동성은 조건부 분산-공분산 행렬 \mathbf{H}_t 을 의미하며 \mathbf{H}_t 에 대한 모형을 설정함으로써 다변량 수익률간의 동적인 관계(dynamic relationship)를 모형화 할 수 있다. \mathbf{H}_t 에 대한 모형에는 단변량-GARCH 모형의 확장 형태인 EWMA 모형, DVEC 모형 및 BEKK 모형 등이 있다 (Tsay, 2010). 단순 확장 모형은 차원 k 가 증가함에 따라 모수가 급격하게 많아진다는 단점이 있다. 조건부 상관계수가 시점에 관계없이 일정하다고 가정하여 모수의 개수를 줄인 CCC 모형과 CCC 모형의 일반화인 DCC 모형 등도 있다. DCC 모형은 시간에 의존하는 조건부 상관계수 행렬을 고려한 모형이며 자세한 내용은 Tsay (2010)를 참고하면 좋을 것이다.

2.1. Exponentially weighted moving average(EWMA) 모형

EWMA 모형은 지수평활법을 이용하여 조건부 분산-공분산 행렬을 추정하는 방법이며 다음과 같이 구할 수 있다.

$$\mathbf{H}_t = (1 - \lambda)\mathbf{a}_{t-1}\mathbf{a}_{t-1}^T + \lambda\mathbf{H}_{t-1},$$

여기서 λ 는 평활 상수이며 0과 1사이 값을 갖는다. \mathbf{a}_t 에 대한 분포를 가정할 경우, λ 를 추정할 수 있으며 일반적으로 λ 값은 RiskMetrics 에서 사용하는 0.94를 이용한다. 위의 식 보면, 바로 전 시점의 변동성에 더 많은 가중치를 주는 것을 알 수 있다.

2.2. BEKK 모형

조건부 분산-공분산 행렬인 H_t 가 양정치 행렬이어야 한다는 제약조건을 성립시키기 위한 Baba-Engle-Kraft-Kroner(BEKK)의 간편 형태인 BEKK(1,1) 모형은 다음과 같다.

$$H_t = CC^T + A(a_{t-1}a_{t-1}^T)A^T + BH_{t-1}B^T,$$

여기서 C 은 하삼각행렬이고, A 와 B 는 $k \times k$ 행렬이다. CC^T 가 양정치이면 H_t 는 항상 양정치가 된다.

2.3. Constant conditional correlation(CCC) 모형

모수의 개수를 줄이기 위한 조건부 상관계수를 상수로 고정시킨 CCC 모형은 다음과 같다.

$$H_t = D_t R D_t = \left(\rho_{ij} \sqrt{h_{iit} h_{jjt}} \right),$$

$$D_t = \text{diag} \left(h_{11t}^{\frac{1}{2}}, \dots, h_{kk}^{\frac{1}{2}} \right), \quad R = (\rho_{ij}),$$

여기서 h_{iit} ($i = 1, \dots, k$)는 i 번째 수익률의 t 시점에서의 조건부 분산을 의미하며 단변량 GARCH로 모형화된다. 또한, R 은 상관계수 행렬로 시간에 따라 불변(time invariant)이다. CCC 모형은 추정의 어려움을 극복한 장점을 가지고 있지만 시간에 따라 변화하는 상관계수의 특성을 고려하지 못한다는 단점이 있다 (Choi 등, 2009).

2.4. Dynamic conditional correlation(DCC) 모형

Engle (2002)은 시간에 의존하는 조건부 상관계수 행렬을 고려하여 다음과 같은 CCC 모형의 일반화된 형태인 DCC 모형을 제안하였다.

$$H_t = D_t R_t D_t,$$

$$D_t = \text{diag} \left(h_{11t}^{\frac{1}{2}}, \dots, h_{kk}^{\frac{1}{2}} \right),$$

$$R_t = J_t Q_t J_t,$$

여기서 h_{iit} 는 CCC 모형과 같이 단변량 GARCH로 모형화되며 $J_t = \text{diag}(q_{11,t}^{-1/2}, \dots, q_{kk,t}^{-1/2})$ 는 대각행렬이다. $Q_t = (q_{ij,t})$ 는 $k \times k$ 양정치 행렬이며 다음을 만족한다.

$$Q_t = (1 - \theta_1 - \theta_2) \bar{Q} + \theta_1 \epsilon_{t-1} \epsilon_{t-1}^T + \theta_2 Q_{t-1}.$$

ϵ_t 는 $\epsilon_{i,t} = a_{i,t} / \sqrt{\sigma_{i,t}}$ 를 원소로 갖는 표준화된 오차 벡터이며 \bar{Q} 는 ϵ_t 의 비조건부 분산-공분산 행렬이다. θ_1 과 θ_2 는 $0 < \theta_1 + \theta_2 < 1$ 을 만족하는 음이 아닌 모수이다 (Tsay, 2010).

3. 실현변동성(realized volatility; RV)

빅데이터 처리 능력의 향상 등으로 고빈도 자료를 이용한 분석이 가능해짐에 따라 기존의 일별 종가로부터 일별 변동성을 추정하는 방법 이외에 고빈도 자료를 활용하여 변동성을 계산하는 방법이 연구되고 있다. Andersen과 Bollerslev (1997)는 5분 단위의 고빈도 자료를 이용하여 실현변동성을 추정하는 방법을 소개하였다. 실현변동성이란 고빈도 자료의 일중 수익률(intra-day return)로 부터 계산한 일중변동성이며 RV로 나타낸다 (Oh and Shin, 2012; Yoon and Hwang, 2015).

3.1. 실현변동성

k 개의 로그 수익률로 구성된 벡터 $\mathbf{r}_t = (r_{1t}, \dots, r_{kt})^T$ 를 시점 t 의 일간 로그 수익률(daily log return) 벡터라고 하자. 일중 로그 수익률(intra-day log return) 벡터는 t 일 중 일정한 간격으로 n 개가 조사되었다고 가정하여 $\mathbf{r}_{t,i} = (r_{1t,i}, \dots, r_{kt,i})^T$, $i = 1, \dots, n$ 로 나타내며 t 일의 i 번째 관측시점의 로그수익률을 의미한다. 국내 주식 시장의 경우 1분 단위로 관측되었다면 9시부터 15시까지 해당하여 $n = 360$ 이 된다. t 일의 일간 로그 수익률 \mathbf{r}_t 는 n 개의 일중 로그 수익률 벡터의 합으로 다음과 같이 나타낼 수 있다.

$$\mathbf{r}_t = \sum_{i=1}^n \mathbf{r}_{t,i}.$$

$\mathbf{r}_t = (r_{1t}, \dots, r_{kt})^T$ 의 조건부 분산-공분산 행렬은 다음과 같이 계산된다 (Tsay, 2010).

$$\text{Var}(\mathbf{r}_t | F_{t-1}) = \sum_{i=1}^n \text{Var}(\mathbf{r}_{t,i} | F_{t-1}) + 2 \sum_{i < j}^n \text{Cov}[\mathbf{r}_{t,i}, \mathbf{r}_{t,j} | F_{t-1}].$$

F_{t-1} 은 $t-1$ 시점까지 포함된 정보 집합을 나타낸다. 일중 로그 수익률 $\mathbf{r}_{t,i} = (r_{1t,i}, \dots, r_{kt,i})^T$ 이 백색 잡음과정이라 가정하면 조건부 분산-공분산 행렬은 다음과 같다.

$$\text{Var}(\mathbf{r}_t | F_{t-1}) = n \text{Var}(\mathbf{r}_{t,i}) = n \left[E(\mathbf{r}_{t,i} \mathbf{r}_{t,i}^T) - E(\mathbf{r}_{t,i}) E(\mathbf{r}_{t,i})^T \right],$$

여기서 $E(\mathbf{r}_{t,i}) = \mathbf{0}$ 으로 가정하면 일간 로그 수익률 \mathbf{r}_t 의 실현변동성 RV_t 은 다음과 같이 정의할 수 있다.

$$\text{RV}_t = \sum_{i=1}^n \mathbf{r}_{t,i} \mathbf{r}_{t,i}^T.$$

Andersen 등 (2003)은 시장미시구조 잡음이 없다는 가정 하에 수익률의 제곱합으로 계산된 실현변동성 RV_t 는 n 이 무한대로 커짐에 따라 변동성의 일치추정량이 됨을 보인바 있다. 그러나 이것은 수익률이 무상관을 갖는다는 가정 하에 성립하는 것이다. 실제로, 고빈도 자료는 비동시적거래, 주문가-체결가 불일치, 틈한 거래 등의 시장미시구조 잡음으로 인해 자기상관(autocorrelation)구조를 가지게 되어 실현변동성에 편향(bias)이 발생하게 한다 (cf. Oh와 Shin, 2012). 이러한 편향을 줄이기 위해 5분 단위 데이터 사용이 권장된다 (Xiao, 2013).

3.2. 편향 수정 실현변동성

1분, 1초와 같이 5분 단위 이상의 고빈도 자료의 사용이 가능해짐에 따라 데이터의 정보손실을 막으면서 편향을 제거하기 위한 방법으로 Zhou (1996)는 1차 자기상관을 고려한 실현변동성을 제안하였다. 일중 로그 수익률 $\mathbf{r}_{t,i} = (r_{1t,i}, \dots, r_{kt,i})^T$ 이 MA(1) 구조를 따르면 \mathbf{r}_t 의 조건부 분산-공분산 행렬은 다음과 같다. 여기서 $E(\mathbf{r}_t) = E(\sum_{i=1}^n \mathbf{r}_{t,i}) = \mathbf{0}$ 이라고 가정한다.

$$\begin{aligned} \text{Var}(\mathbf{r}_t | F_{t-1}) &= \text{Cov} \left(\sum_{i=1}^n \mathbf{r}_{t,i}, \sum_{i=1}^n \mathbf{r}_{t,i} | F_{t-1} \right) = E \left[\left(\sum_{i=1}^n \mathbf{r}_{t,i} \right) \left(\sum_{i=1}^n \mathbf{r}_{t,i} \right)^T \middle| F_{t-1} \right] \\ &= \sum_{i=1}^n E(\mathbf{r}_{t,i} \mathbf{r}_{t,i}^T) + \sum_{i=2}^n E(\mathbf{r}_{t,i} \mathbf{r}_{t,i-1}^T) + \sum_{i=1}^{n-1} E(\mathbf{r}_{t,i} \mathbf{r}_{t,i+1}^T) \\ &= \sum_{i=1}^n \mathbf{r}_{t,i} \mathbf{r}_{t,i}^T + \frac{n}{n-1} \sum_{i=2}^n \mathbf{r}_{t,i} \mathbf{r}_{t,i-1}^T + \frac{n}{n-1} \sum_{i=1}^{n-1} \mathbf{r}_{t,i} \mathbf{r}_{t,i+1}^T. \end{aligned}$$

따라서 MA(1) 구조를 고려한 편향 수정 실현변동성(RV_{AC})은 다음과 같이 정의된다.

$$RV_{AC} = \sum_{i=1}^n \mathbf{r}_{t,i} \mathbf{r}_{t,i}^T + \frac{n}{n-1} \sum_{i=2}^n \mathbf{r}_{t,i} \mathbf{r}_{t,i-1}^T + \frac{n}{n-1} \sum_{i=1}^{n-1} \mathbf{r}_{t,i} \mathbf{r}_{t,i+1}^T.$$

Hansen과 Lunde (2006)은 위의 추정치가 실현변동성의 편향을 제거함을 보였다. 하지만 위 추정치는 일중 로그 수익률이 음의 자기 상관을 가지게 될 경우 변동성이 음의 값을 가질 수 있다는 단점을 가진다 (cf. Oh와 Shin, 2012). 이러한 단점을 해결하기 위해 Hansen과 Lunde (2006)는 Barlett 커널을 사용하여 음이 아닌(nonnegative) 실현변동성을 제안하였다. Hansen과 Lunde (2006)이 제안한 일중 로그 수익률이 1차 자기상관을 가지며 양의 변동성만을 갖는 RV_{AC}은 다음과 같다.

$$RV_{AC} = \sum_{i=1}^n \mathbf{r}_{t,i} \mathbf{r}_{t,i}^T + \frac{n}{n-1} \sum_{i=1}^{n-1} \mathbf{r}_{t,i} \mathbf{r}_{t,i+1}^T.$$

4. 정준상관 및 VaR

4.1. 정준상관분석(canonical correlation analysis; CCA)

정준상관분석은 여러 변수들을 두 변수집단인 종속변수집단과 독립변수집단으로 분류한 뒤, 두 변수집단 사이의 상관구조를 가장 잘 설명하는 선형결합을 찾는 방법이다. 정준상관분석을 통해 두 집단 간의 관계를 살펴볼 수 있다. 본 논문에서는 정준상관분석을 이용하여 실현변동성과 다변량-GARCH 모형 간의 관계를 파악하고자 한다. 종속변수집단을 실현변동성으로 하고 독립변수집단을 다변량-GARCH 모형(EWMA, BEKK, CCC, DCC 모형 등)으로 하여 두 집단 간 상관관계를 가장 잘 설명하는 선형결합을 찾고 정준상관계수를 계산한다. 또한 실현변동성과 각 다변량-GARCH 모형 간의 정준상관계수를 비교한다. 본 절에서 소개하는 정준상관분석의 기본적 개념과 수식은 Seong (1997)을 참고하여 정리하였다.

종속변수 집단은 p 개 확률변수로 구성된 확률벡터 $Y = (y_1, \dots, y_p)^T$ 로 표시하고, q 개 확률변수로 구성된 독립변수 집단은 확률벡터 $X = (x_1, \dots, x_q)^T$ 로 표시하자. 각 변수 집단에서 p 차원 계수벡터 $a = (a_1, \dots, a_p)^T$ 와 q 차원 계수벡터 $b = (b_1, \dots, b_q)^T$ 를 이용한 선형결합 W 와 V 를 다음과 같이 생각해보자.

$$W = a^T Y = a_1 y_1 + \dots + a_p y_p,$$

$$V = b^T X = b_1 x_1 + \dots + b_q x_q.$$

정준상관분석은 확률변수 W 와 V 의 상관계수 $\rho(W, V)$ 를 최대화시킬 수 있는 계수벡터 a 와 b 를 찾는 것이다. 여기서 계수벡터 a 와 b 를 정준계수벡터라 하며 W 와 V 를 정준변수라 한다. 정준변수는 단위 분산을 가지며 두 정준변수사이의 상관계수가 정준상관계수이다.

4.2. Value at risk(VaR)

VaR은 ‘주어진 신뢰수준(confidence level)하에서 목표기간(target horizon)동안 정상적인 시장(normal market)을 전제로 할 때 발생 가능한 최대 손실 금액(maximum loss)’로 정의 된다 (Choi 등, 2009; Kim과 Lee, 2005). l 기간 동안의 보유 금융자산의 가치변동을 $\Delta V(l)$ 라고 할 때 신뢰수준 $100 \times (1 - \alpha)\%$ 에서의 VaR를 다음과 같이 정의한다.

$$\Pr[\Delta V(l) > \text{VaR}] = 1 - \alpha.$$

즉, 신뢰수준 $100 \times (1 - \alpha)\%$ 에서 포트폴리오의 가치 변동이 VaR보다 클 확률은 $(1 - \alpha)$ 이다.

4.2.1. VaR 측정 정규분포를 가정하는 모수적인 방법을 이용하여 VaR을 측정하도록 한다. 수익률 r_t 에 대해 $r_{t+1}|F_t \sim N(\hat{r}_t(1), \hat{\sigma}_t^2(1))$ 와 같은 조건부 분포를 가정하면 백분위수를 이용하여 VaR를 다음과 같이 계산할 수 있다. 이때, F_t 는 t 시점까지의 정보 집합이며 $\hat{r}_t(1)$ 은 수익률의 1시차 조건부 예측치, $\hat{\sigma}_t^2(1)$ 은 수익률의 분산에 대한 1시차 조건부 예측치이다. 유의 수준 α 는 5%와 1%를 사용하는 것이 일반적이다. k 개의 수익률로 이루어진 포트폴리오의 신뢰수준 $100 \times (1 - \alpha)\%$ 에서의 다변량 VaR는 다음과 같이 정의된다 (Hwang 등, 2009).

$$\text{Multivariate-VaR} = \mathbf{w}^T \hat{r}_t(1) - z_\alpha \sqrt{\mathbf{w}^T \hat{\mathbf{H}}_t(1) \mathbf{w}},$$

여기서 \mathbf{w} 는 포트폴리오에서 각 주식 종목들의 투자가중치 벡터를 의미한다. z_α 은 표준정규분포의 $100 \times (1 - \alpha)\%$ 백분위수이며 $\hat{\mathbf{H}}_t(1)$ 은 수익률 벡터에 대한 1시차 조건부 분산-공분산 행렬이다. 본 논문에서는 $\hat{\mathbf{H}}_t(1)$ 의 값을 앞서 소개한 다변량-GARCH 모델을 통해 VaR를 계산한다.

4.2.2. 사후검정(back-testing) VaR 모형의 정확성을 검증하는 방법에는 실패율을 이용한 방법이 있다 (Hwang 등, 2009). 실패율(failure rate)는 검증기간 중에 실제 자산의 수익률이 예측된 VaR을 초과하는 비율을 의미한다. 실패율 계산을 위해 다음과 같은 지시변수가 고려된다.

$$x_t = I \left\{ \mathbf{w}_t^T \mathbf{r}_t < \widehat{\text{VaR}}_t \right\}.$$

시점 t 에서 포트폴리오의 실제 손실이 예측된 VaR보다 크면 $x_t = 1$ 이고 예측된 VaR보다 작으면 $x_t = 0$ 이다. 전체 T 기간 동안 예측된 VaR을 초과하는 총 실패발생 횟수 X_T 는 다음과 같다.

$$X_T = \sum_{t=1}^T x_t.$$

5. 자료 분석 예제

5.1. 자료설명

분석에 사용한 자료는 KOSPI의 시가총액 상위 종목인 삼성전자, 현대차의 거래가격 데이터이다. 2010년 1월 2일부터 2016년 9월 2일까지의 총 1,653개의 일별 자료를 사용하였고 고빈도 자료는 같은 기간 동안 매일의 장내거래시간 9시부터 15시까지의 1분, 5분 단위로 측정된(고빈도) 자료를 이용하였다. 1분 단위 자료의 경우, 일중 수익률의 개수가 $n = 360$ 이어야 하지만 마감 전 10분 동안의 동시호가(유가증권 매매거래 시 동시에 접수된 호가 또는 시간의 선후가 분명하지 않은 호가)로 인하여 하루에 관측된 수익률의 개수는 $n = 350$ 이다. 마찬가지로 5분 단위의 자료도 동시호가로 인해 일중 수익률이 70개 관측되었다.

5.2. 다변량-GARCH 모형 및 실현변동성 적합

두 종목으로 이루어진 벡터시계열 평균 벡터가 $\mathbf{0}$ 이고 자기상관을 보이지 않으므로 $\mathbf{r}_t = \mathbf{a}_t = \mathbf{H}_t^{1/2} \mathbf{e}_t$ 이다. Table 5.1은 삼성전자와 현대차의 수익률에 대해 EWMA, BEKK(1, 1), CCC(1, 1), DCC(1, 1) 모형을 적합 시킨 결과이다.

Table 5.1. Multivariate-GARCH parameter estimates

Model	Parameter estimates
EWMA	$\hat{\lambda} = 0.94$
BEKK(1, 1)	$\begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -1.239 & 1.588 \end{bmatrix} \begin{bmatrix} 0 & -1.239 \\ 0 & 1.588 \end{bmatrix}$ $+ \begin{bmatrix} -0.144 & 0.115 \\ 0.197 & 0.074 \end{bmatrix} \begin{bmatrix} a_{1,t-1}^2 & a_{1,t-1}a_{2,t-1} \\ a_{2,t-1}a_{1,t-1} & a_{2,t-1}^2 \end{bmatrix} \begin{bmatrix} -0.144 & 0.197 \\ 0.115 & 0.074 \end{bmatrix}$ $+ \begin{bmatrix} 0.113 & 0.789 \\ 0.526 & 0.501 \end{bmatrix} \begin{bmatrix} \sigma_{11,t-1} & \sigma_{12,t-1} \\ \sigma_{21,t-1} & \sigma_{22,t-1} \end{bmatrix} \begin{bmatrix} 0.113 & 0.526 \\ 0.789 & 0.501 \end{bmatrix}$
CCC(1, 1)	$h_{11,t} = 0.062 + 0.022a_{1,t-1}^2 + 0.956h_{11,t-1}^2$ $h_{22,t} = 0.159 + 0.036a_{2,t-1}^2 + 0.924h_{22,t-1}^2$ $\hat{R} = \begin{bmatrix} 1 & 0.301 \\ 0.301 & 1 \end{bmatrix}$
DCC(1, 1)	$h_{11,t} = 0.047 + 0.021a_{1,t-1}^2 + 0.962h_{11,t-1}^2$ $h_{22,t} = 0.185 + 0.038a_{2,t-1}^2 + 0.915h_{22,t-1}^2$ $Q_t = (1 - 0.018 - 0.968) \begin{bmatrix} 1 & 0.294 \\ 0.294 & 1 \end{bmatrix} + 0.018\epsilon_{t-1}\epsilon_{t-1}^T + 0.968Q_{t-1}$

EWMA = exponential weighted moving average; CCC = constant conditional correlation; DCC = dynamic conditional correlation.



(a) 5분 단위 실현변동성(RV)



(b) 1분 단위 편향 수정 실현변동성(RV_{AC})

Figure 5.1. Realized volatility.

Figure 5.1은 이변량 수익률벡터를 이용하여 5분 단위의 RV와 1분 단위 RV_{AC}을 나타낸 시도표이다. 그림을 통해 5분 단위의 RV 보다 1분 단위의 RV_{AC}가 삼성전자와 현대차의 변동성을 대체적으로 더 크게 추정하며 상관계수는 5분 단위의 실현변동성이 더 크게 추정하는 것을 알 수 있다.

Table 5.2. First canonical correlation coefficient

X	Y	
	RV(5 min)	RV _{AC} ^(1 min)
EWMA	0.4884	0.5336
BEKK(1, 1)	0.4534	0.5321
CCC(1, 1)	0.4601	0.5085
DCC(1, 1)	0.4599	0.5035

RV = realized volatility; RV_{AC} = bias-corrected RV; EWMA = exponential weighted moving average; CCC = constant conditional correlation; DCC = dynamic conditional correlation.

Table 5.3. Canonical variables

MGARCH	First pair of canonical variables
EWMA	$\hat{W}_1 = 0.9575h_{11,t}^{*(RV_{AC})} + 0.0900h_{22,t}^{*(RV_{AC})}$
	$\hat{V}_1 = 0.5518h_{11,t}^{*(EWMA)} + 0.5850h_{22,t}^{*(EWMA)}$
BEKK(1, 1)	$\hat{W}_1 = 1.0373h_{11,t}^{*(RV_{AC})} - 0.0938h_{22,t}^{*(RV_{AC})}$
	$\hat{V}_1 = 0.0936h_{11,t}^{*(BEKK)} + 0.9126h_{22,t}^{*(BEKK)}$
CCC(1, 1)	$\hat{W}_1 = 0.9579h_{11,t}^{*(RV_{AC})} + 0.0892h_{22,t}^{*(RV_{AC})}$
	$\hat{V}_1 = 0.4949h_{11,t}^{*(CCC)} + 0.6564h_{22,t}^{*(CCC)}$
DCC(1, 1)	$\hat{W}_1 = 0.9506h_{11,t}^{*(RV_{AC})} + 0.1035h_{22,t}^{*(RV_{AC})}$
	$\hat{V}_1 = 0.4835h_{11,t}^{*(DCC)} + 0.6793h_{22,t}^{*(DCC)}$

MGARCH = Multivariate-GARCH; EWMA = exponential weighted moving average; CCC = constant conditional correlation; DCC = dynamic conditional correlation; RV = realized volatility.

5.3. 모형 비교

Table 5.2는 두 가지 실현변동성과 다변량-GARCH 모형들 간에 각각 정준상관분석을 시행하고 얻은 제 1정준상관계수를 나타낸 표이다. 모든 다변량-GARCH 모형에 대해 5분 단위 RV보다 1분 단위 RV_{AC}와의 제 1정준상관계수가 크므로 다변량 GARCH 모형들이 5분 단위의 실현변동성보다 1분 단위의 편향 수정 실현변동성과 더 큰 상관구조를 가진다고 할 수 있다. 5분 단위 실현변동성과 상관관계가 큰 다변량 GARCH 모형은 EWMA, CCC(1, 1), DCC(1, 1), BEKK(1, 1) 순이며 1분 단위 편향 수정 실현변동성은 EWMA, BEKK(1, 1), CCC(1, 1), DCC(1, 1) 순으로 상관관계가 크다.

다변량-GARCH 모형을 5분 단위 RV보다는 상관관계가 더 큰 1분 단위 RV_{AC}와 비교하는 것이 보다 정확한 비교가 가능하다고 판단하여 1분 단위 RV_{AC}와 다변량-GARCH 모형들의 정준변수를 Table 5.3에 나타냈다. 1분 단위 RV_{AC}를 종속변수집단으로 하고 다변량-GARCH로 추정된 변동성을 독립변수집단으로 하는 경우, 표본정준변수 \hat{V}_1 은 BEKK(1, 1)을 제외한 다변량-GARCH 모형들에서 삼성전자 수익률의 변동성과 현대차 수익률의 변동성에 비슷한 가중치를 두고 있다. 또한 모든 다변량-GARCH 모형에 대한 표본정준변수 \hat{W}_1 는 삼성전자의 1분 단위 RV_{AC}에 영향을 많이 받고 현대차의 1분 단위 RV_{AC}에는 영향을 거의 받지 않고 있다.

Table 5.4는 VaR을 계산하고 사후검증을 수행한 결과이다. 보유기간은 1일, 신뢰수준은 95%와 99%이며 삼성전자와 현대차가 동일한 가중치를 가진다는 가정 하에 VaR(%)를 계산하였다. 2015년 9월 1일부터 2016년 9월 2일까지 250개의 관측치를 표본 외 기간으로 설정하여 사후검증을 실시하

Table 5.4. Canonical variables

Model	α	
	0.05	0.01
RV	0.0281	0.0080
RV _{AC}	0.0040	0.0000
EWMA	0.0361	0.0080
BEKK(1, 1)	0.0321	0.0120
CCC(1, 1)	0.0443	0.0120
DCC(1, 1)	0.0370	0.0123

RV = realized volatility; RV_{AC} = bias-corrected RV; EWMA = exponential weighted moving average; CCC = constant conditional correlation; DCC = dynamic conditional correlation.

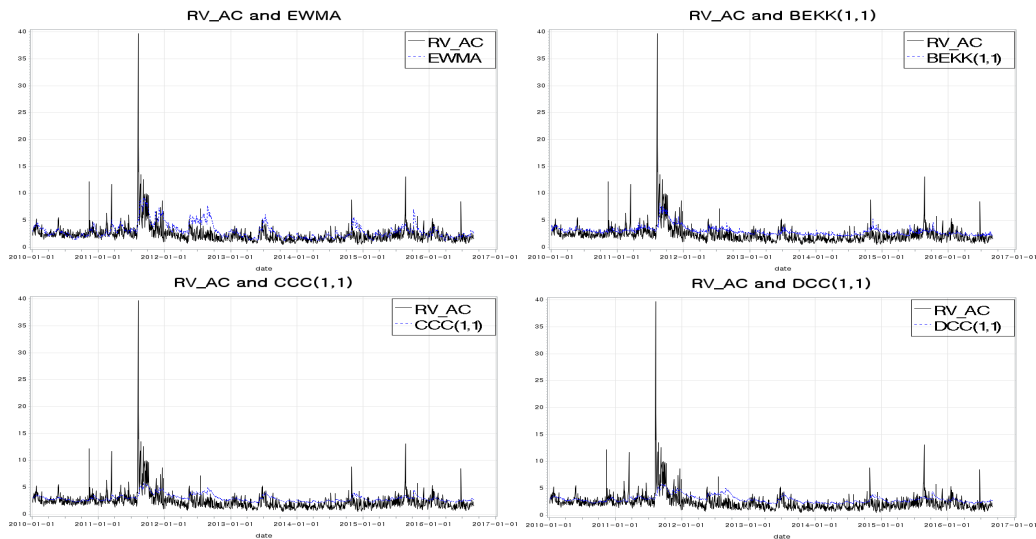


Figure 5.2. Various volatility computations for Samsung electronics.

였다. 실패율을 살펴보면 5분 단위 RV와 1분 단위 RV_{AC}가 다변량-GARCH 모형들에 비해 실패율이 작은 것을 알 수 있으며 RV보다는 RV_{AC}의 실패율이 작다. 이는 RV보다는 RV_{AC}로 구한 VaR이 위험을 정확하게 평가한다는 것을 의미한다. 유의수준 5%에서는 BEKK(1, 1), EWMA, DCC(1, 1), CCC(1, 1) 순으로 실패율이 작으며 유의수준 1%에서는 EWMA, BEKK(1, 1), CCC(1, 1), DCC(1, 1) 순으로 위험을 정확하게 평가한다고 할 수 있다.

Figures 5.2-5.4는 삼성전자와 현대차의 1분 단위 편향 수정 실현변동성과 다변량-GARCH 모형으로 추정된 변동성을 나타낸 시도표이다. 앞서 정준상관분석과 VaR분석 결과에서 CCC(1, 1), DCC(1, 1) 보다는 EWMA, BEKK(1, 1) 모형이 실현변동성과 비슷한 움직임을 가진다는 결과를 그림을 통해서도 확인 할 수 있다.

6. 결론 및 논의

본 연구에서는 고빈도 자료를 이용한 실현변동성을 계산하여 참값으로 간주한 후, 다변량 GARCH 모

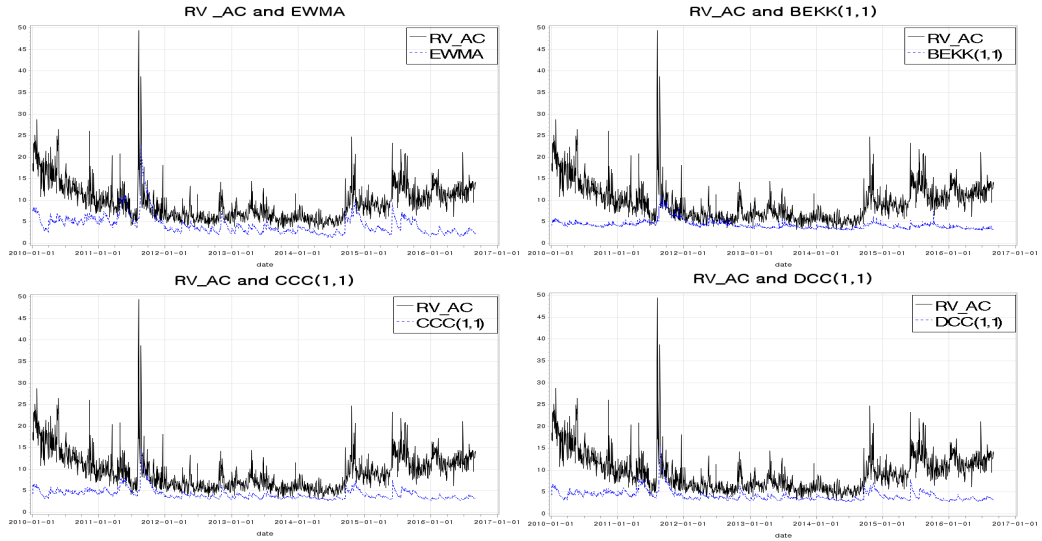


Figure 5.3. Various volatility computations for Hyundai motor.

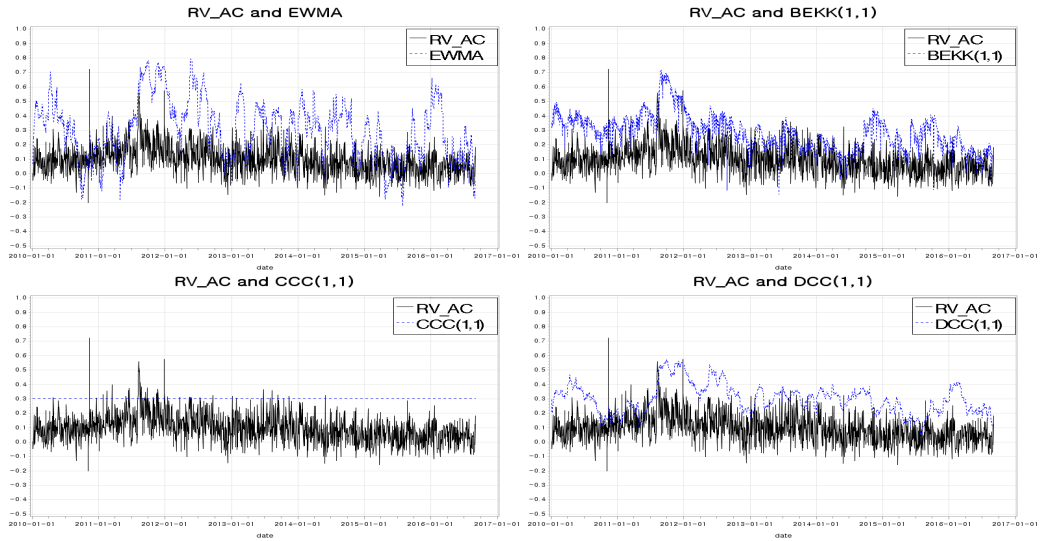


Figure 5.4. Volatility correlations between Samsung electronics and Hyundai motor.

형 선택에 대하여 논의하였다. 고빈도 자료로는 일반적으로 권장되는 5분 단위 자료와 더불어 5분 보다 더 추출간격이 짧은 1분 단위 자료를 사용하였다. 1분 단위 자료의 경우 시장미시구조의 잡음으로 인해 편향이 발생하여 RV_{AC} 을 계산하였다. 모형비교에는 실현변동성과의 상관구조를 고려한 정준상관 분석을 이용하였고 리스크 관리 측면에서 VaR분석 및 실패율을 고려하였다. 삼성전자와 현대차의 거래가격 데이터를 이용하여 이변량 실현변동성과 이변량 GARCH 모형에 적합 시킨 결과, EWMA 모형

과 BEKK(1, 1) 모형이 다른 다변량-GARCH 모형에 비해 실현변동성과 비슷한 움직임을 보이고 리스크 측면에서도 뒤처지지 않았다. 하지만 EWMA와 BEKK 모형은 단변량 GARCH의 단순 확장 모형으로 이변량 이상으로 차원이 높아지면 추정해야할 모수가 많아진다는 단점 있기 때문에 더 높은 차원에서는 다른 다변량-GARCH 모형들과의 비교연구도 필요할 것으로 판단된다.

References

- Andersen, T. G. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets, *Journal of Empirical Finance*, **4**, 115–158.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modelling and forecasting realized volatility, *Econometrics*, **71**, 579–625.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307–327.
- Choi, S. M., Hong, S. Y., Choi, M. S., Park, J. A., Baek, J. S., and Hwang, S. Y. (2009). Analysis of multivariate-GARCH via DCC modeling, *Korean Journal of Applied Statistics*, **22**, 995–1005.
- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987–1007.
- Engle, R. F. (2002). Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business and Economic Statistics*, **20**, 339–350.
- Hansen, P. R. and Lunde, A. (2006). Realized variance and market microstructure noise, *Journal of Business & Economic Statistics*, **24**, 127–161.
- Hwang, S. Y., Choi, M. S., and Do, J. D. (2009). Assessments for multivariate-GARCH models using back-testing : case study, *Korean Journal of Applied Statistics*, **22**, 261–270.
- Kim, H. and Lee, M. (2005). *Econometrics and Financial Time Series*, Kyungmunsa, Seoul.
- Oh, R. and Shin, D. W. (2012). Market microstructure noise and optimal sampling frequencies for the realized variances of stock prices of four leading Korean companies, *Korean Journal of Applied Statistics*, **25**, 15–27.
- Seong, W. H. (1997). *Applied Multivariate Analysis: Theory, Methods, SAS Application*, Tamjin, Seoul.
- Tsay, R. S. (2010). *Analysis of Financial Time Series* (3rd ed), Wiley, New York.
- Xiao, L. (2013). Realized volatility forecasting: empirical evidence from stock market indices and exchange rates, *Applied Financial Economics*, **23**, 57–69.
- Yoon, J. E. and Hwang, S. Y. (2015). Volatility computations for financial time series : high frequency and hybrid method, *Korean Journal of Applied Statistics*, **28**, 1163–1170.
- Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates, *Journal of Business & Economic Statistics*, **14**, 45–52.

다변량 고빈도 금융시계열의 변동성 분석

이근주^a · 황선영^{a,1}

^a숙명여자대학교 통계학과

(2016년 12월 23일 접수, 2017년 1월 18일 수정, 2017년 1월 18일 채택)

요약

본 논문은 다변량 변동성을 다루고 있다. 최근 들어 활발하게 연구가 되고 있는 고빈도(high frequency)자료에 기초한 변동성 측정방법인 실현변동성을 계산하고 기존의 다변량 GARCH 모형과 비교분석하였다. 정준상관분석과 VaR분석을 이용하여 실현변동성과 다양한 다변량 GARCH 모형을 비교하였으며 최근 6년 동안의 삼성전자/현대차 거래 가격 고빈도 데이터를 이용하여 실증분석을 실시하였다.

주요용어: 고빈도 자료, 실현변동성, 다변량 GARCH

본 연구는 한국연구재단의 기초연구과제 사업의 지원을 받았습니다(과제번호: 2015-057031).

¹교신저자: (04310) 서울특별시 용산구 청파로47길 100, 숙명여자대학교 통계학과.

E-mail: shwang@sookmyung.ac.kr