응용통계연구 (2011)
24(4), 597–607
DOI: http://dx.doi.org/10.5351/KJAS.2011.24.4.597

Value at Risk의 사후검증을 통한 다변량 시계열자료의 차원축소 방법의 비교: 사례분석

이대수[1] · 송성주[2]

[1]고려대학교 통계학과, [2]고려대학교 통계학과

(2011년 5월 접수, 2011년 7월 채택)

요 약

금융자산에의 투자에서 리스크 관리의 중요성이 부각되면서 리스크를 측정할 수 있는 도구로서 Value at Risk (VaR)가 널리 각광을 받고 있다. Value at Risk는 주어진 신뢰수준에서 목표기간 동안 발생 가능한 최대손실로 정의되는데 몇 가지 한계점이 있지만 비교적 간단하게 계산되고 이해될 수 있다는 장점이 있어 리스크 측정 및 관리의 기본적인 측도로 이용되고 있다. 그러나 포트폴리오에 포함되는 자산의 숫자가 많아지는 경우 VaR을 계산하는 데에 필수적인 변동성 추정이 매우 어려워지게 된다. 이때 차원축소의 방법을 생각할 수 있는데, 전통적인 인자분석은 시계열자료에 적합한 방법이 아니기 때문에 직접 적용할 수 없고 자료의 자기상관성을 제거하는 방법이 선행되어야 한다. 본 논문에서는 인자분석의 확장 형태인 시계열인자분석을 활용하여 시계열자료의 차원축소과정을 간결하게 하는 방법을 제시하고, 시계열인자분석으로 차원을 축소할 때 기존의 방법을 사용하는 것과 어떠한 차이가 있는지를 실제 금융자료를 이용한 VaR의 사후검증을 통해 분석하였다.

주요용어: 인자분석, 시계열인자분석, VaR(Value at Risk), DCCGARCH, CCCGARCH, 차원축소.

1. 서론

금융자료를 분석할 때 다루게 되는 자료는 흔히 다변량 시계열자료이다. 때에 따라서는 많은 금융상품을 한꺼번에 고려하여 포트폴리오를 구성하게 되므로 다루어야 하는 변수의 개수가 매우 많아지게 된다. 또한 금융자료의 분석에서 매우 중요한 부분으로 변동성을 생각할 수 있다. 차원이 높은 다변량 자료에서 변동성을 계산해야 하는 경우, 각 변수에 대한 분산과 더불어 변수들 사이의 공분산을 추정하여야 하므로 그에 따른 어려움이 발생하게 된다. 변수가 많아질 때 추정해야 할 모수의 개수가 급격하게 증가하는 이러한 문제점은 차원축소의 방법을 통해 일부분 해결할 수 있을 것이다. 차원을 줄이는 것은 어느 정도 정보의 손실을 발생시키지만 많은 모수를 추정함에 따른 오차를 줄이고 해석을 쉽게 하는 장점이 있다.

시계열 자료에서의 차원축소를 연구한 논문으로 Connor (1995)는 다변량 시계열자료의 차원축소를 위한 3가지 모형, 즉 거시적 인자모형(macroeconomic factor model), 근본적 인자모형(fundamental factor model), 그리고 통계적 인자모형(statistical factor model)을 소개하고 성능을 비교하였다. 거시적 인자모형은 물가상승률이나 이자율 등의 거시적 경제변수들을 인자로 사용하는 분석이고, 근본적 인자

이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업이며 (No. 2010-0023191) 제 1저자 이대수의 석사학위논문의 축약본임.
[2]교신저자: (136-701) 서울시 성북구 안암동 5-1, 고려대학교 통계학과, 부교수. E-mail: sjsong@korea.ac.kr

모형은 각 기업의 규모나 배당률, 산업분류, book-to-market ratio 등 기업의 근본 변수들을 이용하는 분석이며, 통계적 인자모형은 최대우도 또는 주성분 분석에 기반을 둔 인자모형이다. Connor (1995)에서는 각 모형을 비교한 결과 근본적 인자모형이나 통계적 인자모형이 거시적 인자모형보다 나은 설명력을 가진다고 보고하였다. 국내연구로는 송유진 등 (2008)에서 다변량 금융시계열자료에 인자분석을 적용하였는데, 통계적 인자 모형과 근본적 인자 모형을 이용하여 다변량 시계열자료의 차원을 축소하고 이로부터 얻은 인자를 이용하여 최종모형들의 MSE(Mean Squared Error)와 MAD(Mean Absolute Deviation)를 구하고, VaR(Value at Risk)을 구하여 비교하였다. 근본적 인자 모형보다 통계적 인자모형으로 차원을 축소하였을 때 예측력이 더 좋아지고 VaR가 보다 정확하게 예측된다는 결과를 얻었다. 그러나 인자분석은 시계열 자료가 아닌 횡단면자료에 적용되도록 개발되었기 때문에 시계열자료에 사용되면 인자분석에 필요한 기본가정이 충족되지 않게 되므로 이를 시계열자료에 바로 적용하는 것은 무리가 있다. 이러한 이유로 황선영 등 (2009)에서는 각 자료의 종속성을 제거하기 위하여 주가자료 각각에 일변량 ARMA모형을 적용하여 나온 잔차에 대해 인자분석을 적용하고, 그 인자점수를 이용한 여러 다변량 GARCH모형을 통해 구해진 VaR를 사후검증을 통하여 비교하였다.

한편, Bauwens 등 (2006)에서는 다변량 GARCH모형 자체에도 모형의 추정방식이나 상황에 따라 다양한 모형이 있음을 조사하였다. 이 다양한 다변량 GARCH모형 중에 최성미 등 (2009)에서는 DCC GARCH(Dynamic Conditional Correlation GARCH)모형을 사용하여 우리나라 주가자료의 VaR를 구하였고, 이를 통해 DCC GARCH모형의 우수성을 다소 확인하였다.

이에 본 논문에서는 시계열자료의 특성을 반영한 시계열인자분석과 전통적인 인자분석에 대해 알아보고 이를 이용하여 인자점수를 계산, DCC GARCH와 CCC GARCH(Constant Conditional Correlation GARCH)모형을 적용하여 VaR를 구하고자 한다. 또한 황선영 등 (2009)에서와 같이 VaR의 사후검증을 통하여 인자분석과 시계열인자분석의 차이점을 살펴 볼 것이다.

2. 차원축소와 다변량 GARCH모형

전통적인 인자분석을 시계열자료에 적용하는 경우 시계열자료의 특성인 자기상관성에 의하여 기본가정이 위배된다. 따라서 기존의 인자분석의 확장으로 자기상관성이 있는 자료에 적당한 시계열인자분석이 제안되었는데 이를 간단히 소개하겠다. 보다 자세한 내용은 Gilbert와 Meijer (2005)를 참조할 수 있다.

t시점의 k차원 시계열자료 $y_t = (y_{1t}, y_{2t}, \ldots, y_{kt})'$는 다음과 같은 인자구조를 갖는데

$$y_t = \alpha_t + B\xi_t + \epsilon_t, \tag{2.1}$$

여기서 α_t는 시간에 따른 y_t의 평균, B는 인자적재행렬, ξ_t는 인자점수벡터이고, y_t는 비정상시계열이나 1차 차분하는 경우 정상이 되는 시계열로 간주한다. 따라서 식 (2.1)은

$$Dy_t \equiv y_t - y_{t-1} = (\alpha_t - \alpha_{t-1}) + B\left(\xi_t - \xi_{t-1}\right) + \left(\epsilon_t - \epsilon_{t-1}\right), \tag{2.2}$$

혹은 차분연산자 D를 이용하여

$$Dy_t = \tau_t + BD\xi_t + D\epsilon_t \tag{2.3}$$

로 쓸 수 있다. 식 (2.3)의 구조는 다시 인자구조를 갖는 방정식으로 볼 수 있다. 이렇게 차분된 자료에 전통적인 인자분석의 추정방식을 적용하면 인자적재행렬인 B를 추정할 수 있게 된다.

다변량 시계열자료의 변동성을 모형화하기 위하여 많은 모형들이 제안되어져 있는데, 이 모형들은 변동성을 추정하는 방식에 따라 구별될 수 있다. 상관계수를 따로 모형화하는 방식으로 전체 분산-공분산행렬을 모형화하는 CCC GARCH모형과 DCC GARCH모형 등이 제안되어 있으며, 이 모형들에 관한 다음의 내용은 최성미 등 (2009)를 참조하였다.

다변량 시계열자료에서 t시점의 k차원 수익률을 $r_t = (r_{1t}, r_{2t}, \ldots, r_{kt})'$라 하는 경우 이를 $\mu_t + a_t$로 모형화한다. 이 때 μ_t는 $t-1$시점까지의 정보가 주어져 있을 때 r_t의 조건부 기댓값이며 $a_t = (a_{1t}, \ldots, a_{kt})'$는 t시점에서의 충격으로서 $a_t = H_t^{1/2} u_t$를 만족한다. H_t는 $k \times k$ 조건부 분산-공분산행렬이며 $u_t = (u_{1t}, \ldots, u_{kt})'$는 평균이 0이고 분산-공분산행렬이 $k \times k$ 차원 단위행렬인 랜덤벡터이다. 변동성을 모형화하는 GARCH모형에서는 수익률의 조건부 표준편차로 이루어진 대각행렬 D_t와 조건부 상관계수행렬 R_t에 대해 $H_t = D_t R_t D_t$로 표현할 수 있다.

다변량 GARCH모형을 추정할 때 Bollerslev (1990)는 조건부 상관계수들을 상수로 고정시켜 모수의 수를 줄인 CCC모형을 제안하였는데, 이 모형에서는 H_t를 다음과 같이 정의한다.

$$H_t = D_t R_t D_t = \left(\rho_{ij}\sqrt{h_{ii,t}h_{jj,t}}\right),$$
$$D_t = \mathrm{diag}\left(h_{11,t}^{\frac{1}{2}}, \ldots, h_{kk,t}^{\frac{1}{2}}\right),$$

여기서 $h_{ii,t}$는 t시점의 i번째 자산의 조건부 분산이며 이는 일변량 GARCH모형을 이용하여 계산할 수 있다. 만약 GARCH$(1,1)$을 따르는 CCC모형이라면 $i = 1, \ldots, k$에 대하여

$$h_{ii,t} = w_i + \alpha_i a_{i,t-1}^2 + \beta_i h_{ii,t-1}$$

과 같이 쓸 수 있다. $R = (\rho_{ij})$는 모든 i에 대해서 $\rho_{ii} = 1$이 성립하는 대칭인 양정치 행렬이며 상수인 조건부 상관계수 ρ_{ij}를 가진다.

CCC GARCH는 조건부 상관계수 행렬이 시간에 따라 변하지 않는다는 가정을 통해 모형을 단순화시켰지만, 이 가정이 실제 금융자료에서 만족되기는 어렵다. 따라서 Engle (2002), Tse와 Tsui (2002)는 시간에 의존하는 조건부 상관계수 행렬로 CCC모형을 일반화하였는데 이 모형이 DCC GARCH이다. Engle (2002)에서 소개된 DCC$(1,1)$에서 H_t는 다음과 같이 표현된다.

$$H_t = D_t R_t D_t = \left(\rho_{ij,t}\sqrt{h_{ii,t}h_{jj,t}}\right),$$
$$D_t = \mathrm{diag}\left(h_{11,t}^{\frac{1}{2}}, \ldots, h_{kk,t}^{\frac{1}{2}}\right),$$
$$R_t = \mathrm{diag}\left(q_{11,t}^{-\frac{1}{2}}, \ldots, q_{kk,t}^{-\frac{1}{2}}\right) Q_t \ \mathrm{diag}\left(q_{11,t}^{-\frac{1}{2}}, \ldots, q_{kk,t}^{-\frac{1}{2}}\right),$$

여기서 $k \times k$의 대칭 양정치 행렬인 $Q_t = (q_{ij,t})$는 다음과 같이 정의되며, 이로 인해 조건부 상관계수 행렬 R_t는 시간에 따라 계속 변화하는 형태를 갖게 된다.

$$Q_t = (1-\alpha-\beta)\bar{Q} + \alpha u_{t-1}u_{t-1}' + \beta Q_{t-1}.$$

$\bar{Q}$는 표준화된 오차항 $u_t = H_t^{-1/2} a_t$에 대한 비조건부 분산행렬이며, α와 β는 $\alpha + \beta < 1$을 만족하는 음이 아닌 값을 가지는 모수이다. 이와 같이 CCC GARCH와 DCC GARCH는 여러 다변량 GARCH 모형들 중에서 추정할 모수의 수가 적으면서도 조건부 상관계수들을 모형화하여 추정이 가능하며, 특히 DCC모형은 상관계수가 시간에 따라 변화할 수 있도록 하여 모형을 더욱 일반화하였다.

3. VaR(Value at Risk)와 사후검증

VaR은 정상적인 시장을 전제로 할 때 주어진 신뢰수준에서 목표기간 동안 발생 가능한 최대손실로 정의될 수 있다. L기간 동안 포트폴리오의 가치변동을 $\Delta V(L)$라고 할 때 신뢰수준 $100(1-\alpha)\%$에서의 VaR은 다음과 같이 나타낼 수 있다.

$$\Pr[\Delta V(L) \leq \mathrm{VaR}] = 1 - \alpha. \tag{3.1}$$

k차원의 시계열자료로 포트폴리오를 구성할 때, $(t+1)$ 시점에서의 포트폴리오 수익률은 가중치 벡터 w와 $(t+1)$ 시점에서의 수익률벡터 r_{t+1}을 이용하여 $w'r_{t+1}$로 나타낼 수 있으며, $r_{t+1} \sim N(\hat{r}_t(1), \hat{H}_t(1))$와 같은 조건부 분포를 이용하여 목표기간 1일에 대한 VaR값을 계산할 수 있다. $\hat{r}_t(1)$은 수익률 벡터에 대한 1시차 조건부 예측치이며 $\hat{H}_t(1)$은 수익률 벡터에 대한 1시차 조건부 분산-공분산 행렬의 예측치이다. 신뢰수준 $100(1-\alpha)\%$에서의 다변량 VaR은 z_α가 정규분포의 $100(1-\alpha)$ 백분위수일 때 다음과 같이 계산된다.

$$\text{다변량 VaR} = w'\hat{r}_t(1) - z_\alpha \sqrt{w'\hat{H}_t(1)w}, \tag{3.2}$$

여기서 $\hat{H}_t(1)$의 값은 다양한 다변량 GARCH모형으로 예측할 수 있다.

VaR의 정확성을 검증하는 과정을 사후검증이라고 부르는데 가장 간편하게 VaR의 정확성을 검증할 수 있는 방법은 실패율을 이용하는 방법이라고 할 수 있다. 실패율방법은 주어진 모형을 이용하여 예측한 VaR값 보다 실제 손실이 더 커지는 것을 모형의 실패라 보고 검증 기간 동안 그 횟수를 세어 실패율을 측정하는 방법이다. 실패율계산을 위해 다음과 같은 지시변수를 정의하도록 하겠다.

$$x_t = I\left\{ w_t' r_t < \widehat{\mathrm{VaR}}_t \right\}, \tag{3.3}$$

여기서 $\widehat{\mathrm{VaR}}_t$는 모형에 의해 추정된 VaR의 값이다. 특정한 날(t)의 포트폴리오의 실제손실이 예측된 VaR보다 크면 $x_t = 1$, 예측된 VaR보다 작으면 $x_t = 0$이 되며, T일 동안 예측된 VaR를 초과하여 발생하는 총 실패횟수 X_T는 $\sum_{t=1}^{T} x_t$로 쓸 수 있다. VaR측정에 사용된 모형이 적절하다면 $\{x_t, t = 1, 2, \ldots, T\}$값들은 서로 독립적이며, 주어진 신뢰수준이 $1-\alpha$라면 각 x_t가 평균이 α인 베르누이 분포를 따르게 된다. 따라서 X_T는 기댓값이 $T\alpha$이고 분산은 $T\alpha(1-\alpha)$인 이항분포를 따르게 된다. 이러한 점을 이용하여 실패율을 검정하는 Kupiec의 검정은 다음과 같은 가설을 검정한다.

$$\begin{aligned} H_0 &: \alpha = \alpha_0. \\ H_a &: \alpha \neq \alpha_0. \end{aligned} \tag{3.4}$$

α_0는 목표하고 있는 VaR의 실패율이며 실제의 실패율이 목표하고 있는 VaR의 신뢰수준과 같은지를 귀무가설 하에서 근사적으로 $\chi^2(1)$을 따르는 다음과 같은 통계량을 사용하여 검정하게 된다.

$$-2\ln\left\{ \frac{\alpha^{X_T}(1-\alpha)^{T-X_T}}{(X_T/T)^{X_T}(1-X_T/T)^{T-X_T}} \right\}. \tag{3.5}$$

또 다른 검정방법으로 Christoffersen의 검정은 x_t들의 현 시점과 전 시점을 비교하여 다음과 같은 방식으로 x_t들의 독립성을 검정한다. $t = 1, 2, \ldots, T$에서 관측치를 얻을 때, x_{t-1}의 값이 i이면서 x_t의 값이 j가 되는 모든 날의 합을 n_{ij}라고 정의하자 ($i, j = 0$ 또는 1). 즉, 모든 $t = 1, 2, \ldots, T$에 대해 $x_{t-1} = 0$이고 $x_t = 0$인 경우는 n_{00}에 1을, $x_{t-1} = 0$이고 $x_t = 1$인 경우는 n_{01}에 1을, $x_{t-1} = 1$이고

$x_t = 0$인 경우는 n_{10}에 1을, $x_{t-1} = 1$이고 $x_t = 1$인 경우는 n_{11}에 1을 더하면 n_{ij}의 값을 얻을 수 있다. 이 값들을 이용, $\pi_{ij} = n_{ij}/\sum_j n_{ij}$를 계산하여 식 (3.6)의 가설을 검정한다.

$$\begin{aligned} H_0 &: x_t\text{들은 서로 독립이다.} \\ H_a &: x_t\text{들은 독립이 아니다.} \end{aligned} \tag{3.6}$$

사용되는 검정통계량은

$$2\ln\left\{(1-\pi_{01})^{n_{00}}\,\pi_{01}^{n_{01}}\,(1-\pi_{11})^{n_{10}}\,\pi_{11}^{n_{11}}\right\} - 2\ln\left\{\alpha^{X_T}\,(1-\alpha)^{T-X_T}\right\}$$

이며, 이는 귀무가설 하에서 근사적으로 $\chi^2(1)$의 분포를 따른다. 사후검증에 대한 더 자세한 내용은 Kupiec (1995)과 Christoffersen과 Pellitier (2004)를 참조할 수 있다.

4. 실증 분석

국내 13개 대기업 주가의 로그차분수익률을 전통적인 인자분석과 시계열인자분석으로 차원을 축소하여 구해진 요인점수에 다변량 GARCH모형을 적용, VaR을 구하여 비교해 보았다. 사용한 프로그램은 통계 프로그램 R의 tsfa, ccgarch, vars패키지 등이다. 분석에 사용된 자료는 국내 13개 대기업의 주가로서 2004년 1월14일부터 2008년 8월 13일까지의 총 1137일의 로그차분수익률(%)인데, P_t를 t시점에서의 주가라 했을 때 로그차분 수익률은 $R_t = \ln(P_t/P_{t-1}) \times 100$으로 정의된다. 이 자료는 황선영 등 (2009)에서 다루어진 자료와 거의 같은 자료이다. 황선영 등 (2009)에서 사용된 자료와 같이 다음의 13개 대기업으로 분석을 진행하였다.

- 은행: 신한지주, 우리금융, 기업은행, 외환은행.
- 건설: 현대건설, 대림산업, GS건설, 대우건설.
- IT: 삼성전자, LG전자, 하이닉스, LS산전, 엔씨소프트.

4.1. 차원축소모형의 적합

위의 자료를 이용하여 인자분석과 시계열인자분석을 시행하여 비교하고자 하였다. 인자분석의 경우, 시계열자료의 특성인 자기상관성에 의하여 분석을 바로 적용하지 못하기 때문에 각 주가의 로그차분수익률을 적당한 ARMA모형에 적합한 후 그 잔차에 인자분석을 시행하였는데, 인자의 수는 그림 4.1의 스크리도표를 활용하여 결정하였다. 인자의 수를 k라고 하였을 때 k에 해당하는 그래프상의 값은 k번째 특성치까지의 누적 비율을 나타낸다. 따라서 인자의 수가 3인 경우에서 4인 경우로의 누적비의 증가가 별로 없기에 인자의 수를 3개로 정하였다.

또한 시계열인자분석을 시행하였는데, 시계열인자분석은 기본가정에 의하여 자기상관성이 있어도 되기 때문에 ARMA모형의 잔차가 아닌 직접 로그차분수익률 자료에 시행하여 추정하였다. 표 4.3은 시계열인자분석의 적합도통계량들로서 확증적 인자분석에서 사용하는 기준인 RMSEA(root mean squared error of approximation), GFI(goodness-of-fit index), AGFI(adjusted goodness-of-fit index)를 사용하여 인자의 개수를 정하였다. 적합도통계량에 대한 자세한 내용은 김기영과 강현철 (2001)을 참조할 수 있다.

확증적 인자분석에서 RMSEA는 0.05보다 작은 경우, GFI와 AGFI는 0.9보다 높은 경우 모형이 잘 적합된 것으로 판단한다. 이 기준을 활용하면 3개의 인자 이후로는 모형들이 모두 적합한 것으로 나타나고 있으므로 가장 간단한 모형인 3개의 인자를 인자의 수로 결정하였다.

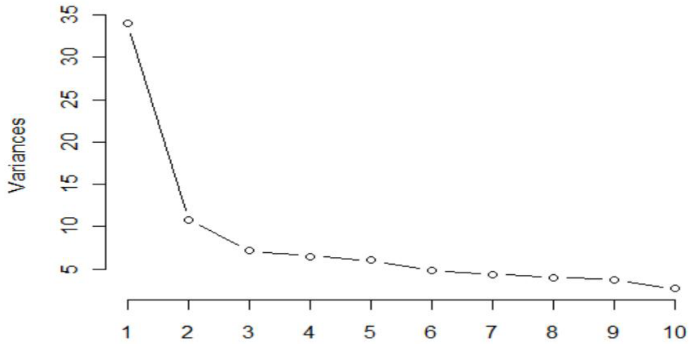

그림 4.1. 인자분석의 스크리도표

표 4.1. 인자분석의 적재행렬

	회전된 인자패턴		
	요인 1	요인 2	요인 3
신한지주	0.221	**0.643**	0.328
우리은행	0.240	**0.705**	0.226
기업은행	0.238	**0.696**	0.262
외환은행	0.226	**0.570**	0.224
현대건설	**0.725**	0.147	0.107
대림산업	**0.581**	0.195	0.233
GS건설	**0.639**	0.202	0.175
대우건설	**0.621**	0.209	0.111
삼성전자	0.139	0.331	**0.747**
LG전자	0.201	0.176	**0.583**
하이닉스	0.130	0.201	**0.544**
LS산전	**0.390**	0.185	0.293
엔씨소프트	0.226	0.178	**0.292**

표 4.1과 표 4.2는 인자분석과 시계열인자분석에서의 인자적재행렬들이다. 인자적재행렬의 내용을 보면 시계열인자분석과 전통적인 인자분석 모두 실제기업들의 분류와 대부분 일치하는 것을 볼 수 있는데, 이로 인해 3개의 인자가 적합한 것으로 보인다. 한 가지 특이한 것은 LS산전의 경우 다른 기업들과 달리 요인점수가 실제 분류와 일치하지 않게 나타난다는 점이다. 이것은 LS산전의 기업특성이 다른 IT관련 기업들과는 다르기 때문으로 보이며, 건설주나 IT주로 확연히 구분하기 어려운 그 특성을 고려할 때 오히려 IT 요인의 점수가 높게 나오지 않는 것이 더 적절할 것이라고 생각된다.

실제기업의 분류와 인자분석의 적재행렬을 통하여 요인 1은 건설주, 요인 2는 은행주, 요인 3은 IT주로 명명하도록 하겠다. 시계열 인자분석에서는 순서가 바뀌어 요인 1이 은행주, 요인 2가 건설주, 요인 3이 IT주가 된다.

4.2. 다변량 GARCH 모형적합

CCC GARCH와 DCC GARCH는 각각의 조건부분산을 일변량 GARCH모형으로, 시계열자료들간의 관계는 상관행렬을 이용하여 모형화하는 모형들이다. 이 두 모형을 적합한 후 잔차들에 자기상관성이 남아 있는지를 검정하여 모형이 잘 적합 되었는지 확인하였다.

표 **4.2.** 시계열인자분석의 인자적재행렬

	회전된 인자패턴		
	요인 1	요인 2	요인 3
신한지주	**0.680**	−0.028	0.111
우리은행	**0.759**	0.037	−0.041
기업은행	**0.723**	−0.006	0.046
외환은행	**0.691**	0.019	−0.057
현대건설	−0.058	**0.811**	−0.078
대림산업	0.047	**0.488**	0.151
GS건설	0.047	**0.572**	0.107
대우건설	0.050	**0.664**	−0.052
삼성전자	0.136	−0.061	**0.743**
LG전자	−0.074	0.055	**0.678**
하이닉스	0.027	−0.007	**0.559**
LS산전	0.098	**0.312**	0.220
엔씨소프트	0.111	0.134	**0.259**

표 **4.3.** 시계열인자분석 적합통계량

인자수	통계량		
	RMSEA	GFI	AGFI
2	0.078	0.939	0.896
3	0.048	0.979	0.955
4	0.042	0.987	0.963
5	0.031	0.993	0.974

표 **4.4.** 인자분석 요인점수의 다변량 GARCH모형 적합 전과 후의 잔차제곱의 다변량포트만투 통계량 (괄호 안은 p-value)

모형적합전	$Q(6) = 298(0.0001)$ $Q(12) = 516(0.0001)$	$Q(18) = 621(0.0001)$ $Q(24) = 721(0.0001)$ $Q(30) = 878(0.0001)$
CCC GARCH 적합후	$Q(6) = 37.845(0.953)$ $Q(12) = 104.209(0.585)$	$Q(18) = 148.651(0.766)$ $Q(24) = 193.63(0.86)$ $Q(30) = 259.932(0.659)$
DCC GARCH 적합후	$Q(6) = 37.066(0.937)$ $Q(12) = 106.545(0.522)$	$Q(18) = 149.494(0.741)$ $Q(24) = 195.426(0.839)$ $Q(30) = 262.31(0.620)$

표 4.4와 표 4.5에 모형적합전의 다변량 포트만투 검정통계량과 적합후의 통계량을 표기하였다. 통계량은 6, 12, 18, 24, 30시차까지의 값을 계산하였으며 괄호안의 숫자들은 p-value를 나타낸다. 검정 결과 인자분석 요인점수에서는 모형적합 이후 자기상관성이 존재하지 않아 모형적합이 잘 된 것을 확인할 수 있었다. 하지만 시계열인자분석의 요인점수에서는 CCC GARCH 모형적합 이후의 잔차에 자기상관성이 존재하여 모형적합이 좋지 않음을 확인할 수 있었다.

4.3. 사후검증

위에서 적합시킨 다변량 GARCH모형들로 분산-공분산행렬을 예측하여 VaR을 구해보았다. 보유기간은 1일, 신뢰수준은 95%이며, 포트폴리오를 구성하는 산업에 투자하는 비중이 다르다는 것을 요인에 가

표 4.5. 시계열인자분석 요인점수의 다변량 GARCH모형 적합 전과 후의 잔차제곱의 다변량포트만투 통계량 (괄호 안은 p-value)

모형적합전	$Q(6) = 485(0.0001)$ $Q(12) = 801(0.0001)$	$Q(18) = 958(0.0001)$ $Q(24) = 1101(0.0001)$ $Q(30) = 1258(0.0001)$
CCC GARCH 적합후	$Q(6) = 77.9511(0.181)$ $Q(12) = 153(0.0028)$	$Q(18) = 191.861(0.0544)$ $Q(24) = 254.4338(0.0374)$ $Q(30) = 312.5374(0.0383)$
DCC GARCH 적합후	$Q(6) = 67.364(0.105)$ $Q(12) = 136.135(0.0348)$	$Q(18) = 174.220(0.242)$ $Q(24) = 233.077(0.202)$ $Q(30) = 287.304(0.224)$

표 4.6. 인자분석 요인점수의 CCC GARCH 실패율

가중치(은행, 건설, IT)	실패율	UC	IND
1/2, 1/4, 1/4	0.0521	0.0225	1.3278
1/4, 1/2, 1/4	0.0608	0.5366	1.5733
1/4, 1/4, 1/2	0.0739	2.4290	2.7293
1/3, 1/3, 1/3	0.0739	2.4290	2.4177
7/10, 2/10, 1/10	0.0698	0.5366	1.3412
2/10, 7/10, 1/10	0.1043	11.0430**	3.5370
1/10, 2/10, 7/10	0.0652	1.0274	2.1045

표 4.7. 인자분석 요인점수의 DCC GARCH 실패율

가중치(은행, 건설, IT)	실패율	UC	IND
1/2, 1/4, 1/4	0.0826	4.3390*	2.7500*
1/4, 1/2, 1/4	0.1043	11.0403**	4.3140*
1/4, 1/4, 1/2	0.1043	11.0400**	5.6324*
1/3, 1/3, 1/3	0.1000	9.5009**	4.2907*
7/10, 2/10, 1/10	0.0826	4.3399*	2.4434
2/10, 7/10, 1/10	0.1260	20.0876**	5.7868*
1/10, 2/10, 7/10	0.1000	9.5009**	5.1466*

중치를 달리하는 것으로 나타내었다. 2007년 9월 3일부터 2008년 8월 13일까지의 230개의 관측치를 표본 외 기간으로 설정하여 사후검증을 실시하였다.

표 4.6∼표 4.9에서 UC는 Kupiec 통계량의 값이고 IND는 Christoffersen 통계량의 값이다. 검정통계량에서 *표시는 p-value가 0.01이상이고 0.05미만의 값을 갖는 경우를, **표시는 p-value가 0.01이하의 값을 갖는 경우를 나타낸다. 표에서 나타난 실패율을 비교해보면 인자분석 요인점수의 CCC GARCH 실패율과 시계열인자분석 요인점수의 DCC GARCH 실패율이 비슷하다는 것을 확인 할 수 있고 인자분석의 요인점수에 DCC GARCH모형을 적용하여 구한 VaR은 목표한 실패율을 맞추고 있지 못하는 것을 알 수 있다. 또한 시계열인자분석의 요인점수에 CCC GARCH모형을 적용하여 구한 VaR도 목표하고 있는 실패율을 맞추고 있지 못하는 것을 확인할 수 있다. 목표하고 있는 실패율을 맞추고 있는 방법은 인자분석의 요인점수에 CCC GARCH모형을 적용하여 구한 VaR과 시계열인자분석의 요인점수에 DCC GARCH모형을 적용하여 구한 VaR이다. 이 둘의 실패율이 가중치에 따라 변하는지를 확인하기 위하여 가중치를 Uniform분포에서 발생시켜 100번의 실패율계산의 사후검증을 시행하였다. 다음 표 4.10은 100개의 랜덤한 가중치를 이용한 사후검증 결과 UC통계량과 IND통계량의 기각횟수를 보여

표 **4.8.** 시계열인자분석 요인점수의 CCC GARCH 실패율

가중치(은행, 건설, IT)	실패율	UC	IND
1/2, 1/4, 1/4	0.0739	2.4290	2.4177
1/4, 1/2, 1/4	0.0826	4.3399*	2.7568
1/4, 1/4, 1/2	0.0826	4.3399*	3.4428
1/3, 1/3, 1/3	0.0826	4.3392*	3.0898
7/10, 2/10, 1/10	0.0652	1.0274	1.8333
2/10, 7/10, 1/10	0.0782	3.3242	2.1372
1/10, 2/10, 7/10	0.0895	5.4704*	3.8336

표 **4.9.** 시계열인자분석 요인점수의 DCC GARCH 실패율

가중치(은행, 건설, IT)	실패율	UC	IND
1/2, 1/4, 1/4	0.0521	0.0225	1.1159
1/4, 1/2, 1/4	0.0695	1.6610	1.8428
1/4, 1/4, 1/2	0.0739	2.4290	2.7293
1/3, 1/3, 1/3	0.0696	1.6610	2.1147
7/10, 2/10, 1/10	0.0782	0.0232	1.1105
2/10, 7/10, 1/10	0.0608	0.5366	1.5733
1/10, 2/10, 7/10	0.0826	4.3392*	3.4428

표 **4.10.** 100개의 랜덤한 가중치를 이용한 사후검증결과: UC통계량과 IND통계량 기각횟수

	인자분석 CCC GARCH	시계열인자분석 DCC GARCH
UC통계량 기각횟수	49	1
IND통계량 기각횟수	1	0

준다.

표 4.10은 가중치를 달리하여도 시계열인자분석의 요인점수에 DCC GARCH모형을 적용하여 구한 VaR는 목표하고 있는 실패율을 맞추지 못하는 경우가 적지만 인자분석의 요인점수에 CCC GARCH모형을 적용하여 구한 VaR은 목표하고 있는 실패율을 맞추지 못하는 경우가 49번으로 가중치의 영향에 민감하게 반응하는 것을 보여준다. 또, IND 통계량 기각횟수를 통해 독립성에는 두 방법 모두 문제가 없는 것으로 나타났다.

위의 표 4.10, 표 4.6, 표 4.9를 통하여 인자분석을 시행할 때는 CCC GARCH모형으로, 시계열인자분석을 시행할 때에는 DCC GARCH모형으로 적합하는 것이 VaR의 실패율을 낮출 수 있다는 것을 확인할 수 있었다. 또한 비교하고 있는 4가지의 방법 중에서 시계열인자분석으로 차원축소를 한 후에 DCC GARCH모형으로 모형을 적합하여 VaR를 측정하는 방법이 실패율을 가장 낮출 수 있다는 것을 알 수 있었다.

5. 결론

본 연구에서는 다변량 시계열자료의 차원을 인자분석으로 축소하여 다변량 GARCH모형을 적합하는 경우, VaR의 사후검증을 통해 인자분석과 시계열 인자분석, 또 CCC GARCH와 DCC GARCH의 성능을 비교하였다. 특히, 본 논문에서는 황선영 등 (2009) 등의 기존 연구에서 했듯이 인자분석을 적용하기 전에 시계열자료의 자기상관성을 제거하고 그 잔차에 인자분석을 시행한 다음 다변량 GARCH모형을 적용하는 복잡한 과정을 거치지 않고, 시계열인자분석을 시계열자료에 직접 적용하는 단순한 방법을

사용하였다.

실증분석의 자료는 2004년 1월14일부터 2008년 8월 13일의 로그차분수익률을 사용하였으며, 다변량 GARCH모형을 적합한 후에 표준화된 잔차의 제곱에 포트만투검정을 실시하여 모형의 적합성을 검정하였다. 이 과정에서 시계열인자분석을 적용한 후 CCC GARCH모형을 적용하는 경우에는 잔차의 제곱에 자기상관성이 존재하는 것을 확인하였고, 나머지 모형들은 적절히 적합되는 것을 알 수 있었다.

다변량 GARCH모형들을 사용하여 VaR을 구한 후 사후검증을 한 결과에서는 인자분석의 요인점수에 DCC GARCH모형을 적용하여 구한 VaR은 목표하고 있는 실패율을 맞추고 있지 못하는 것으로 나타나 인자분석의 요인점수에는 DCC GARCH모형은 적합하지 않다는 것을 확인하였으며, 인자분석 CCC GARCH모형과 시계열인자분석 DCC GARCH모형을 통하여 구한 VaR은 목표하고 있는 실패율을 맞추고 있는 것으로 나타났다. 이 두 가지의 방법의 비교로서 가중치를 랜덤하게 주어 두 모형을 비교한 결과 가중치에 따라서 인자분석 CCC GARCH모형은 실패하는 횟수가 49번으로 많지만 시계열인자분석 DCC GARCH모형은 1번의 실패를 보이는 것으로 나타났다. 따라서 시계열인자분석 요인점수에 DCC GARCH모형을 통하여 VaR을 구하는 경우가 4가지 방법 가운데 가장 안정적인 것으로 나타났다.

위의 결과에 비추어 볼 때, 다변량 시계열자료의 차원축소에서는 인자분석보다 시계열인자분석이 사후검증을 통한 관점에서 더 안정적이고 간결한 방법임을 알 수 있었다.

참고문헌

김기영, 강현철 (2001). <LISREL(SIMPLIS)를 이용한 구조방정식모형의 분석>, 자유아카데미.

송유진, 최문선, 황선영 (2008). 차원축소를 통한 다변량 시계열의 변동성분석 및 응용, <한국통계학회 논문집>, **15**, 825–835.

최성미, 홍선영, 최문선, 박진아, 백지선, 황선영 (2009). DCC 모델링을 이용한 다변량-GARCH 모형의 분석 및 응용, <응용통계연구>, **22**, 995–1005.

황선영, 최문선, 도종두 (2009). 사후검증(Back-testing)을통한 다변량 GARCH 모형의 평가: 사례분석, <응용통계연구>, **22**, 261–270.

Bauwens, L., Laurent, S. and Rombouts, J. V. K. (2006). Multivariate GARCH models, A survey, *Journal of Applied Econometrics*, **21**, 79–109.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model, *The Review of Economics and Statistics*, **72**, 498–505.

Christoffersen, P. and Pellitier, D. (2004). Backtesting value-at-risk: A duration-based approach, *Journal of Financial Econometrics*, **2**, 84–108.

Connor, G. (1995). The three type of factor models, A Comparison of their explanatory power, *Financial Analysts Journal*, **51**, 42–46.

Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate GARCH models, *Journal of Business and Economic Statistics*, **20**, 339–350.

Gilbert, P. D. and Meijer, E. (2005). *Time Series Factor Analysis with an Application to Measuring Money*, Research Report 05F10, University of Groningen, SOM Research School.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models, *Journal of Derivatives*, **3**, 73–84.

Tse, Y. K. and Tsui, A. K. C. (2002). A multivariate GARCH model with time-varying correlations, *Journal of Business and Economic Statistics*, **20**, 351–362.

Comparison of Dimension Reduction Methods for Time Series Factor Analysis: A Case Study

Daesu Lee[1] · Seongjoo Song[2]

[1]Department of Statistics, Korea University; [2]Department of Statistics, Korea University

(Received May 2011; accepted July 2011)

Abstract

Value at Risk(VaR) is being widely used as a simple tool for measuring financial risk. Although VaR has a few weak points, it is used as a basic risk measure due to its simplicity and easiness of understanding. However, it becomes very difficult to estimate the volatility of the portfolio (essential to compute its VaR) when the number of assets in the portfolio is large. In this case, we can consider the application of a dimension reduction technique; however, the ordinary factor analysis cannot be applied directly to financial data due to autocorrelation. In this paper, we suggest a dimension reduction method that uses the time-series factor analysis and DCC(Dynamic Conditional Correlation) GARCH model. We also compare the method using time-series factor analysis with the existing method using ordinary factor analysis by backtesting the VaR of real data from the Korean stock market.

Keywords: Factor analysis, time series factor analysis, Value at Risk, DCCGARCH, CCC GARCH, dimension reduction.

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (No. 2010-0023191). This research is part of Daesu Lee's (the first author) Master thesis.

[2]Corresponding author: Associate Professor, Department of Statistics, Korea University, Seoul 136-701, Korea. E-mail: sjsong@korea.ac.kr