

A Graphical Improvement in Volatility Analysis for Financial Series

Jeong Won Lee^a · Jae Eun Yoon^a · Sun Young Hwang^{a,1}

^aDepartment of Statistics, Sookmyung Women's University

(Received August 12, 2013; Revised August 26, 2013; Accepted August 26, 2013)

Abstract

News Impact Curves(NIC) developed by Engle and Ng (1993) have been useful for graphically representing the volatilities arising from financial time series. Adding an improvement and refinement to the original NIC, this article proposes so called two dimensional NIC and principal component NIC. We illustrate the methodology via Kosdaq data.

Keywords: News Impact Curve, principal component, volatility.

1. 서론

KOSDAQ, KOSPI, 환율과 같은 금융시계열의 변동성(volatility, 조건부 분산, h_t 로 표현)을 살펴보면 수익률에 따라 값이 변동하는 성질을 볼 수 있으며 과거 시점에서의 수익률에 따라 그 변동성이 달라진다. Engle과 Ng (1993)에 의해 처음 소개된 News Impact Curve(NIC)는 $t - 2$ 이전 시점까지의 정보는 상수로 고정시키고, $t - 1$ 시점의 수익률과 그 다음 시점의 변동성인 h_t 간의 관계를 그래프로 표현한 것이다. 이는 $t - 1$ 시점에서의 수익률이 한 시점 후의 변동성에 얼마나 영향을 주는지 측정할 수 있게 한다.

본 논문에서는 기존의 NIC에서 $t - 2$ 시점의 수익률을 추가한 그래프인 이차원 NIC(2D-NIC)를 소개하고, $t - 1$ 시점에서부터 $t - 5$ 시점까지의 수익률의 선형결합인 주성분 분석(Principal Components Analysis; PCA) 방법을 통하여 새로운 인공축인 C_1 (제1주성분, the 1st principal component)으로 나타낸 그래프(PCA in NIC)를 소개하고자 한다. 주성분 분석은 서로 상관관계가 있는 변수들 사이의 복잡한 구조를 좀 더 간단히 이해하기 쉽게 설명하기 위하여 사용하는 분석기법이다. 이는 변수들의 선형결합을 통하여 변수들이 가지고 있는 전체정보를 최대한 설명할 수 있는 서로 독립적인 새로운 인공변수(artificial variable)들을 유도하여 해석하는 기법이며 (Seong, 1997, Chapter 4) 금융시계열에서는 차원축소 방법으로 널리 이용되고 있다. 시계열 변동성의 특성을 잘 파악하기 위해서는 과거 수익률과 변동성과의 관계가 어떻게 나타나는지 쉽게 볼 수 있어야 하므로 본 논문에서 제시하는 새로운 그래프 기법들은 시각적으로 보기 쉽게 나타내기 위한 실용적인 방법으로서 응용성이 있으리라 판단하며, 국내 금융시계열 자료인 Kosdaq에 적용해 보고자 한다.

¹Corresponding author: Professor, Department of Statistics, Sookmyung Women's University, Yongsan-Gu, Seoul 142-742, Korea. E-mail: shwang@sookmyung.ac.kr

2. GARCH모형과 T-GARCH모형

금융시계열의 변동성(volatility, 조건부 분산, h_t 로 표현)은 변동성 집중(volatility clustering)현상, 비대칭성 현상, 급침(leptokurtic) 등이 있다. 이러한 성질들을 다른 일반 시계열 자료에서 잘 나타나지 않는 것으로 이 같은 특징을 잘 나타내는 모형으로 Engle (1982)은 ARCH(autoregressive conditional heteroscedastic)모형을 제안하였다. 이어서, Bollerslev (1986)는 ARCH모형의 확장으로서 GARCH모형을 다음과 같이 제안하였다.

$$\begin{aligned}\epsilon_t &= \sqrt{h_t}e_t, \\ h_t &= \alpha_0 + \sum_{i=1}^p \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}\end{aligned}$$

여기서, $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0$ 을 의미하고, $\{\epsilon_t\}$ 는 평균이 0이고 분산이 1인 *i.i.d* 과정이다. 위의 모형을 우리는 GARCH(p, q)로 나타낸다. GARCH모형에서 오차항이 제곱의 형태로 나타나 있으므로 현재의 조건부 분산에 대하여 오차항의 부호는 영향을 미치지 않으며 현재의 조건부 분산은 대칭인 형태를 보인다. 이는 변동성에 대한 “leverage effect”를 효과적으로 나타내지 못한다는 단점이 있다. 여기서 “leverage effect”는 비대칭 변동성 효과로서 같은 양이라면 수익률의 하락이 상승보다 더 큰 변동성을 제공한다는 의미이다. Rabemananjara와 Zakoian (1993)은 변동성의 비대칭성을 반영하는 모형, T-GARCH(Threshold GARCH)모형을 제시하였다. T-GARCH(1, 1)모형은 다음과 같다.

$$\begin{aligned}\epsilon_t &= \sqrt{h_t}e_t, \\ h_t &= \alpha_0 + \alpha_{11} (\epsilon_{t-1}^+)^2 + \alpha_{12} (\epsilon_{t-1}^-)^2 + \beta_1 h_{t-1}\end{aligned}$$

여기서, $\alpha_{11} \geq 0, \alpha_{12} \geq 0, \epsilon^+ = \max(\epsilon, 0), \epsilon^- = \max(-\epsilon, 0)$ 이다. 참고로 $\epsilon = \epsilon^+ - \epsilon^-$ 와 $|\epsilon| = \epsilon^+ + \epsilon^-$ 의 성질을 지니고 있다. T-GARCH모형에서 모수 α_{11} 와 α_{12} 가 다른 값을 가짐으로써 비대칭성을 반영할 수 있게 된다. 이 모형은 좋은 정보와 나쁜 정보 각각에 따라 다르게 반응함으로써 “leverage effect”를 설명하기에 효과적이다. 여기서 만일 $\alpha_{11} = \alpha_{12}$ 가 된다면, 위의 T-GARCH(1, 1)모형은 GARCH(1, 1)모형과 같게 된다.

2.1. News Impact Curve

Engle과 Ng (1993)에 의해서 처음 소개된 News Impact Curve(NIC)는 $t - 2$ 시점 이하의 모두 상수로 변환한 후에 t 시점에서의 변동성과 $t - 1$ 시점에서의 수익률간의 관계를 쉽게 보기 위하여 그래프로 표현한 것이다. 여기서 NIC는 2차원 좌표 평면에서 수익률 ϵ_{t-1} 은 가로축에, 변동성 h_t 는 세로축에 나타낸다. 이는 비대칭적 효과(leverage effect)를 쉽게 확인하는데 적합한 도구이다. 좋은 정보는 NIC에서 $\epsilon_{t-1} > 0$ 부분에서 변동성의 기울기를, 나쁜 정보는 $\epsilon_{t-1} < 0$ 부분에서 변동성의 기울기를 붉으로써 이를 시각적으로 쉽게 확인 할 수 있다. 만약 “leverage effect”가 존재 한다면, $\epsilon_{t-1} < 0$ 부분에서의 기울기가 $\epsilon_{t-1} > 0$ 에서의 기울기 보다 더 클 것이다.

GARCH모형에서의 NIC 방정식은 다음과 같다.

$$\begin{aligned}h_t &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \cdot \frac{\alpha_0}{1 - \alpha_1 - \beta_1} \\ &= \alpha_1 \epsilon_{t-1}^2 + \alpha_0 + \frac{\alpha_0 \beta_1}{1 - \alpha_1 - \beta_1}.\end{aligned}$$

T-GARCH모형에서의 NIC 방정식은 다음과 같다.

$$\begin{aligned} h_t &= \alpha_0 + \alpha_{11} (\epsilon_{t-1}^+)^2 + \alpha_{12} (\epsilon_{t-1}^-)^2 + \beta_1 h_{t-1} \\ &= \alpha_0 + \alpha_{11} (\epsilon_{t-1}^+)^2 + \alpha_{12} (\epsilon_{t-1}^-)^2 + \frac{\alpha_0 \beta_1}{1 - \frac{1}{2}(\alpha_{11} + \alpha_{12}) - \beta_1}. \end{aligned}$$

좀 더 일반적인 모형에서의 NIC 방정식은 Choi 등 (2012)을 참고하기 바란다.

2.2. 2D-NIC

Engle과 Ng (1993)이 소개한 NIC는 $t-1$ 시점의 수익률과 t 시점의 변동성 간의 관계를 시각적으로 나타내었다. 이를 통해 시각적으로 비대칭적 효과를 쉽게 확인할 수 있으나, 특정 시점의 변동성과 바로 이전 시점의 수익률 간의 관계만 비교 가능하다는 한계가 있다. 본 연구에서는 이를 보완하기 위하여 기존의 NIC에 $t-2$ 시점의 수익률을 반영하는 하나의 축을 추가하여 3차원의 NIC를 제안하였으며 이를 2D-NIC라 명명하였다. $t-2$ 시점의 수익률을 추가함으로써 h_t , ϵ_{t-1} , ϵ_{t-2} 간의 관계를 한눈에 시각적으로 볼 수 있다. 이는 기존의 NIC와 다르게 $t-2$ 시점에서의 수익률과 t 시점에서의 변동성간의 관계를 볼 수 있다. 2D-NIC에서는 한 시점 전과 두 시점 전까지의 수익률을 고려하여 변동성을 나타내므로 기존의 NIC보다 예측의 신뢰도가 올라간다고 볼 수 있다. GARCH모형에서의 2D-NIC 방정식은 다음과 같이 유도할 수 있다.

$$\begin{aligned} h_t &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} \\ &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 [\alpha_0 + \alpha_1 \epsilon_{t-2}^2 + \beta_1 h_{t-2}] \\ &= (1 + \beta_1) \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_1 \beta_1 \epsilon_{t-2}^2 + \beta_1^2 h_{t-2} \\ \therefore h_t &= (1 + \beta_1) \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_1 \beta_1 \epsilon_{t-2}^2 + \frac{\alpha_0 \beta_1^2}{1 - \alpha_1 - \beta_1}. \end{aligned}$$

GARCH모형의 2D-NIC은 $\epsilon_{t-1} = 0$ 과 $\epsilon_{t-2} = 0$ 에서 3차원 포물선의 중심이고, 양방향으로 대칭인 함수이다.

T-GARCH모형에서의 2D-NIC 방정식은 다음과 같다.

$$\begin{aligned} h_t &= \alpha_0 + \alpha_{11} \epsilon_{t-1}^{+2} + \alpha_{12} \epsilon_{t-1}^{-2} + \beta_1 h_{t-1} \\ &= \alpha_0 + \alpha_{11} \epsilon_{t-1}^{+2} + \alpha_{12} \epsilon_{t-1}^{-2} + \beta_1 [\alpha_0 + \alpha_{11} \epsilon_{t-2}^{+2} + \alpha_{12} \epsilon_{t-2}^{-2} + \beta_1 h_{t-2}] \\ &= (1 + \beta_1) \alpha_0 + \alpha_{11} (\epsilon_{t-1}^{+2} + \beta_1 \epsilon_{t-2}^{+2}) + \alpha_{12} (\epsilon_{t-1}^{-2} + \beta_1 \epsilon_{t-2}^{-2}) + \beta_1^2 h_{t-2} \end{aligned}$$

여기서, h_{t-2} 를 $E(h_{t-2}) = \alpha_0 / \{1 - (1/2)(\alpha_{11} + \alpha_{12}) - \beta_1\}$ 으로 바꾼다.

그러므로 $h_t = (1 + \beta_1) \alpha_0 + \alpha_{11} (\epsilon_{t-1}^{+2} + \beta_1 \epsilon_{t-2}^{+2}) + \alpha_{12} (\epsilon_{t-1}^{-2} + \beta_1 \epsilon_{t-2}^{-2}) + \alpha_0 \beta_1^2 / \{1 - (1/2)(\alpha_{11} + \alpha_{12}) - \beta_1\}$.

T-GARCH모형의 2D-NIC는 $\epsilon_{t-1} = 0$ 과 $\epsilon_{t-2} = 0$ 에서 함수의 최솟값을 갖고, 이를 중심으로 양쪽 방향으로 비대칭적으로 증가한다.

2.3. PCA in NIC

주성분 분석은 변수들 사이에 서로 상관관계가 있는 자료 구조를 변수들의 선형결합으로 이루어진 새로

은 인공변수를 사용하여 나타내려는 분석 방법이다. 주성분 분석의 가장 일반적인 두 가지 목표는 차원 축소와 해석의 용이함이다. 일반적으로 변수가 증가하면 이에 따른 공분산 또는 상관계수는 함께 증가하기 마련이며 이와 동시에 변수들의 해석에도 어려움이 따를 것이다. 이러한 문제점들을 조금이나마 줄이기 위하여 우리는 주성분 분석을 사용한다.

확률벡터(random vector) $X^T = [x_1, x_2, \dots, x_p]$ 가 존재한다고 하자. 이의 공분산 행렬 Σ 은 p 개의 고유값(eigenvalue)을 지니고 있다. 고유값의 크기는 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ 이다. 새로운 인공변수 $C_i (1 \leq i \leq p)$ 에 대하여 $X_j (1 \leq j \leq p)$ 들의 선형 결합 형태를 보면 다음과 같다.

$$\begin{aligned} C_1 &= a_1^T X = a_{11}x_1 + a_{12}x_2 + \dots + a_{1p}x_p \\ C_2 &= a_2^T X = a_{21}x_1 + a_{22}x_2 + \dots + a_{2p}x_p \\ &\vdots \\ C_p &= a_p^T X = a_{p1}x_1 + a_{p2}x_2 + \dots + a_{pp}x_p, \end{aligned}$$

여기서 우리는 다음을 얻을 수 있다.

$$\begin{aligned} \text{Var}(C_i) &= a_i^T \Sigma a_i, \quad i = 1, 2, \dots, p, \\ \text{Cov}(C_i, C_j) &= a_i^T \Sigma a_j, \quad i, j = 1, 2, \dots, p. \end{aligned} \quad (2.1)$$

다음의 주성분 C_i 들은 식 (2.1)의 분산들이 가능한 가장 크고, 상관관계가 존재하지 않는 X_i 들의 선형 결합이다. 즉, 제1주성분은 식 (2.1)에서 가장 큰 값을 갖는 변수이고 이를 우리는 C_1 이라 표현한다. 주성분 분석을 시행하면 종종 분석 전의 변수들 사이에서는 보이지 않았던 연관 관계가 분석 후에 보여서 해석이 용이할 때가 있다. 우리는 이러한 내용을 토대로 GARCH모형과 T-GARCH모형에서 찾을 수 없었던 수익률과 변동성 사이의 관계를 주성분 분석을 통하여 찾아보고자 한다. PCA in NIC는 기존의 NIC와 비슷하지만, 가로축을 $t-1$ 시점에서의 수익률 대신에 $t-1$ 시점부터 $t-k$ 시점까지의 수익률들로 이루어진 제1주성분 C_1 을 이용한다는 점에서 차이가 있다.

GARCH(1,1)모형에서의 PCA in NIC를 보기 위하여 다음과 같이 식을 변형하였다. 편의를 위하여 수익률 변수에 대한 주성분 분석은 $t-1$ 시점부터 $t-5$ 시점까지이다.

$$\begin{aligned} h_t &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 h_{t-1} \\ &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 (\alpha_0 + \alpha_1 \epsilon_{t-2}^2 + \beta_1 h_{t-2}) \\ &= (1 + \beta_1) \alpha_0 + (\epsilon_{t-1}^2 + \beta_1 \epsilon_{t-2}^2) \alpha_1 + \beta_1^2 h_{t-2} \\ &\vdots \\ &= (1 + \beta_1 + \dots + \beta_4) \alpha_0 + (\epsilon_{t-1}^2 + \beta_1 \epsilon_{t-2}^2 + \dots + \beta_1^4 \epsilon_{t-5}^2) \alpha_1 + \beta_1^5 h_{t-5} \end{aligned} \quad (2.2)$$

여기서, h_{t-5} 를 $E(h_{t-5})$ 인 상수로 대체하고 $(1 + \beta_1 + \dots + \beta_4) \alpha_0 + \beta_1^5 E(h_{t-5})$ 를 상수 ψ_0 으로 놓는다. $\epsilon_{t-1}^2, \epsilon_{t-2}^2, \dots, \epsilon_{t-5}^2$ 를 성분으로 하는 제1주성분을 C_1 이라 하면 식 (2.2)은 다음과 같이 된다.

$$h_t \approx \psi_0 + C_1.$$

즉, GARCH(1,1)모형에서의 PCA in NIC은 $\epsilon_{t-1}^2, \epsilon_{t-2}^2, \dots, \epsilon_{t-5}^2$ 을 성분으로 사용하면 된다. T-GARCH(1,1)모형에서의 PCA in NIC를 보기 위하여 다음과 같이 식을 변형하였다. 위와 동일하

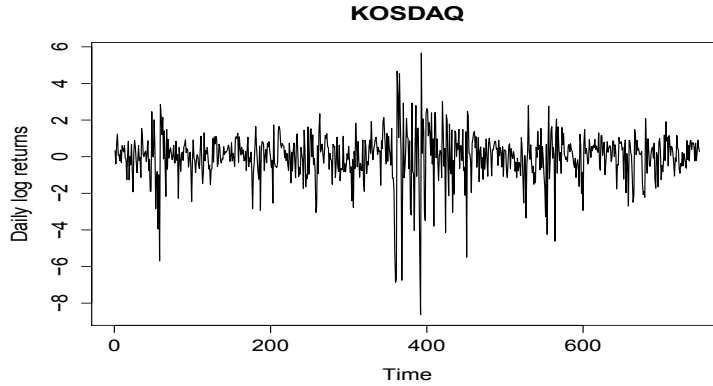


Figure 3.1. Time series plot of returns in KOSDAQ

계 수익률 변수에 대한 주성분 분석은 $t - 1$ 시점부터 $t - 5$ 시점까지이다.

$$\begin{aligned}
 h_t &= \alpha_0 + \alpha_{11}\epsilon_{t-1}^{+2} + \alpha_{12}\epsilon_{t-1}^{-2} + \beta_1 h_{t-1} \\
 &= \alpha_0 + \alpha_{11}\epsilon_{t-1}^{+2} + \alpha_{12}\epsilon_{t-1}^{-2} + \beta_1 [\alpha_0 + \alpha_{11}\epsilon_{t-2}^{+2} + \alpha_{12}\epsilon_{t-2}^{-2} + \beta_1 h_{t-2}] \\
 &= \alpha_0 + \beta_1 \alpha_0 + \alpha_{11} (\epsilon_{t-1}^{+2} + \beta_1 \epsilon_{t-2}^{+2}) + \alpha_{12} (\epsilon_{t-1}^{-2} + \beta_1 \epsilon_{t-2}^{-2}) + \beta_1^2 h_{t-2} \\
 &= (1 + \beta_1)\alpha_0 + \alpha_{11} (\epsilon_{t-1}^{+2} + \beta_1 \epsilon_{t-2}^{+2}) + \alpha_{12} (\epsilon_{t-1}^{-2} + \beta_1 \epsilon_{t-2}^{-2}) + \beta_1^2 [\alpha_0 + \alpha_{11}\epsilon_{t-3}^{+2} + \alpha_{12}\epsilon_{t-3}^{-2} + \beta_1 h_{t-3}] \\
 &\vdots \\
 &= (1 + \beta_1 + \dots + \beta_4)\alpha_0 + \alpha_{11} (\epsilon_{t-1}^{+2} + \beta_1 \epsilon_{t-2}^{+2} + \dots + \beta_1^4 \epsilon_{t-5}^{+2}) \\
 &\quad + \alpha_{12} (\epsilon_{t-1}^{-2} + \beta_1 \epsilon_{t-2}^{-2} + \dots + \beta_1^4 \epsilon_{t-5}^{-2}) + \beta_1^5 h_{t-5}, \tag{2.3}
 \end{aligned}$$

여기서 h_{t-5} 를 $E(h_{t-5})$ 인 상수로 대체하고 $(1 + \beta_1 + \dots + \beta_4)\alpha_0 + \beta_1^5 E(h_{t-5})$ 를 상수 ψ_0 으로 놓는다. $\epsilon_{t-1}^{+2}, \epsilon_{t-2}^{+2}, \epsilon_{t-3}^{+2}, \epsilon_{t-4}^{+2}, \epsilon_{t-5}^{+2}, \epsilon_{t-1}^{-2}, \epsilon_{t-2}^{-2}, \epsilon_{t-3}^{-2}, \epsilon_{t-4}^{-2}, \epsilon_{t-5}^{-2}$ 를 성분으로 하는 제1주성분을 C_1 이라 하면 식 (2.3)은 다음과 같이 된다. $h_t \approx \psi_0 + C_1$. 즉, T-GARCH(1,1)모형에서의 PCA in NIC은 $\epsilon_{t-1}^{+2}, \epsilon_{t-2}^{+2}, \epsilon_{t-3}^{+2}, \epsilon_{t-4}^{+2}, \epsilon_{t-5}^{+2}, \epsilon_{t-1}^{-2}, \epsilon_{t-2}^{-2}, \epsilon_{t-3}^{-2}, \epsilon_{t-4}^{-2}, \epsilon_{t-5}^{-2}$ 을 성분으로 사용하면 된다.

3. 국내금융시계열에의 응용 : KOSDAQ

본 절에서는 2010년 3월 2일부터 2013년 3월 2일까지의 국내 금융시계열 자료 KOSDAQ을 이용하여 앞 절에서 소개한 GARCH(1,1)모형과 T-GARCH(1,1)모형에 적합 시켜 보았다. 그 후 NIC, 2D-NIC, 그리고 PCA in NIC를 적용시켜 시각적으로 나타난 그림을 바탕으로 결과를 비교 분석해 보았다. KOSDAQ 자료의 수익률을 모형에 적합시키기 위하여 로그 차분 후 백분율로 변형한(p_t 로 나타내기로 하자) 후 분석에 이용하였다. 이 자료는 일별시계열로 p_t 로 변형된 자료의 시도표는 Figure 3.1과 같다. 2D-NIC와 PCA in NIC를 효과적으로 비교하기 위하여 GARCH(1,1)모형과 T-GARCH(1,1)모형에 대한 기존의 News Impact Curve를 제시하였다. 모든 분석 과정은 수익률이 (과거 수익률에 조건부적으로) 표준정규분포를 따른다고 가정한 경우와 t 분포를 따른다고 가정한 경우 모두 같으므로 표준정규분포를 따르는 경우에 대해서만 자세히 기술하고 나머지 경우에 대해서는 분석 결과를 요약하여 제시하였다.

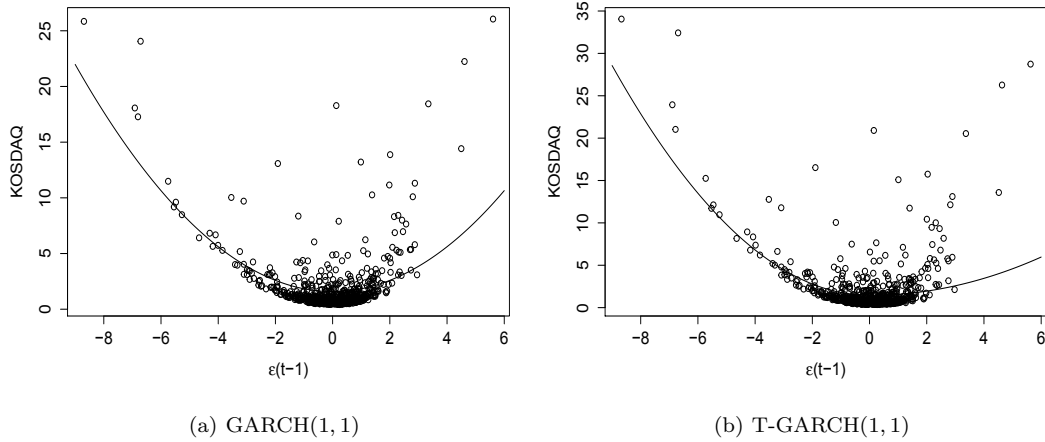


Figure 3.2. NIC of KOSDAQ

3.1. 사례분석 1 : 수익률 ϵ_t 이 조건부 $N(0, 1)$ 인 경우의 KOSDAQ 자료

2010년 3월 2일부터 2013년 3월 2일까지 749개의 KOSDAQ 수익률 데이터를 이용하여 GARCH(1, 1) 모형과 T-GARCH(1, 1)모형을 적합시켜 보았다. 모형의 추정식은 아래와 같다.

$$\begin{aligned}
 \text{GARCH ; } p_t &= 0.054202 + \epsilon_t \\
 h_t &= 0.107282 + 0.251797\epsilon_{t-1}^2 + 0.697492h_{t-1}, \\
 \text{T-GARCH ; } p_t &= 0.03238 + \epsilon_t \\
 h_t &= 0.085188 + 0.125574\epsilon_{t-1}^{+2} + 0.33481\epsilon_{t-1}^{-2} + 0.724472h_{t-1}.
 \end{aligned}$$

GARCH모형은 $\epsilon_{t-1} = 0$ 을 중심으로 부호에 상관없이 대칭인 모형이다. 그러나 T-GARCH모형은 이와 다르게 비대칭 모형이다. 이는 Figure 3.2(a)와 Figure 3.2(b)의 News Impact Curve에서도 확인할 수 있다. 추정식에서 T-GARCH모형은 ϵ_{t-1}^- 에 의한 변동성이 ϵ_{t-1}^+ 에 대한 변동성 보다 3배 더 큰 leverage effect를 갖는 것을 볼 수 있다. 즉, 나쁜 정보가 좋은 정보보다 변동성에 더 크게 영향을 주고 있다. 이러한 성질은 Figure 3.2(b)에서도 음의 부분의 기울기가 양의 부분의 기울기 보다 더 가파름을 통해 확연히 볼 수 있다.

2D-NIC에 대한 추정식은 다음과 같다.

$$\begin{aligned}
 \text{GARCH ; } h_t &= 0.251797\epsilon_{t-1}^2 + 0.175626\epsilon_{t-2}^2 + 1.211318, \\
 \text{T-GARCH ; } h_t &= \begin{cases} 0.125574\epsilon_{t-1}^{+2} + 0.090975\epsilon_{t-2}^{+2} + 1.133135, & \text{if } \epsilon_{t-1} > 0 \text{ and } \epsilon_{t-2} > 0, \\ 0.125574\epsilon_{t-1}^{+2} + 0.242561\epsilon_{t-2}^{-2} + 1.133135, & \text{if } \epsilon_{t-1} > 0 \text{ and } \epsilon_{t-2} < 0, \\ 0.334810\epsilon_{t-1}^{-2} + 0.090975\epsilon_{t-2}^{+2} + 1.133135, & \text{if } \epsilon_{t-1} < 0 \text{ and } \epsilon_{t-2} > 0, \\ 0.334810\epsilon_{t-1}^{-2} + 0.242561\epsilon_{t-2}^{-2} + 1.133135, & \text{if } \epsilon_{t-1} < 0 \text{ and } \epsilon_{t-2} < 0. \end{cases}
 \end{aligned}$$

이를 그래프로 나타내면 다음 Figure 3.3과 같다. Figure 3.3(a)를 보면, $(\epsilon_{t-1}, \epsilon_{t-2}) = (0, 0)$ 을 중심으로 모든 방향에서 대칭인 포물선이다. 이는 $t-2$ 시점이 일정하게 고정되어 있을 때, $t-1$ 시점에서의 수익률은 t 시점에서의 변동성에 부호와 관계없이 같은 영향을 준다. 반대로, $t-1$ 시점이 고정되어 있을

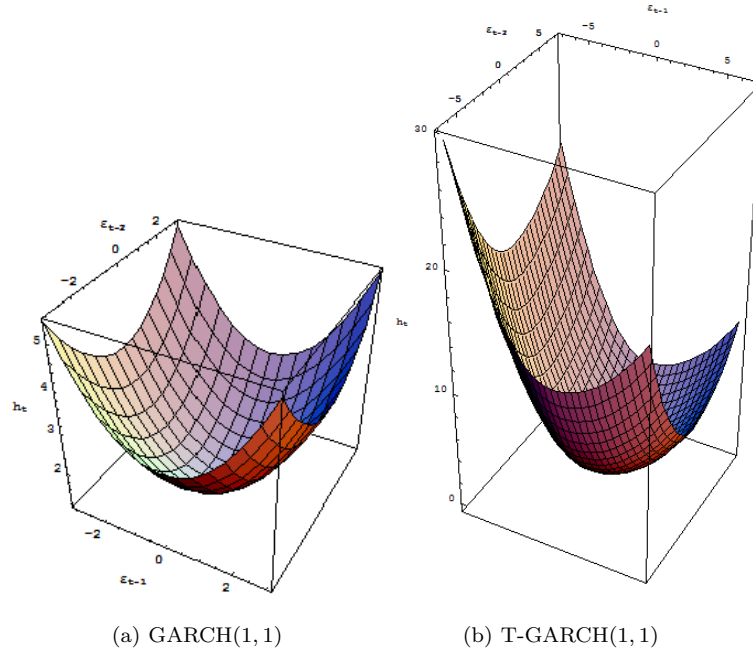


Figure 3.3. 2D-NIC of KOSDAQ (standard normal distribution)

때 $t - 2$ 시점과 t 시점에서의 관계도 위와 같다. 즉, GARCH모형에서는 $t - 1$ 시점의 수익률과 $t - 2$ 시점의 수익률은 부호에 관계없이 t 시점에서의 변동성에 영향을 준다고 볼 수 있다.

T-GARCH모형인 2D-NIC를 보면, $(\epsilon_{t-1}, \epsilon_{t-2}) = (0, 0)$ 에서 t 시점 변동성의 최솟값을 가지고, 모든 방향으로 비대칭적으로 증가함을 Figure 3.3(b)에서 볼 수 있다. 또한, $t - 1$ 시점에서의 수익률과 $t - 2$ 시점에서의 수익률 모두 음의 방향으로 증가할 때 t 시점의 변동성에 가장 크게 영향을 준다. 2D-NIC는 이와 같이 t 시점의 변동성과 $t - 1$ 시점과 $t - 2$ 시점의 수익률의 영향력을 동시에 볼 수 있다는 장점이 있다.

다음으로, KOSDAQ 자료에서 t 시점 이전의 수익률들이 t 시점에서의 변동성에 얼마나 영향이 있는지 보고자 PCA in NIC을 그려보았다. $t - 1$ 시점부터 $t - 5$ 시점까지의 수익률들만 고려하였다. 먼저 GARCH(1,1)모형에서의 PCA in NIC의 성분은 식 (2.2)에 의하여 $\epsilon_{t-1}^2, \epsilon_{t-2}^2, \epsilon_{t-3}^2, \epsilon_{t-4}^2, \epsilon_{t-5}^2$ 이다. 이들의 표본 covariance matrix를 보면, Table 3.1과 같다. 수익률은 정상성을 지니고 있으므로 시차에 관계없이 표본분산은 약 30.9로 일정함을 볼 수 있다. 또한 시차의 간격이 일정한 수익률들 사이의 공분산도 동일함을 확인할 수 있다. 다음의 covariance matrix를 사용하여 5개의 수익률들을 성분으로 하는 주성분분석을 실시한 후, 해당되는 제1주성분, C_1 을 추출하였다. 주성분 분석에 대한 제1주성분 분석의 결과는 다음과 같다.

$$C_1 = 0.39468\epsilon_{t-1}^2 + 0.469547\epsilon_{t-2}^2 + 0.496785\epsilon_{t-3}^2 + 0.46985\epsilon_{t-4}^2 + 0.395221\epsilon_{t-5}^2.$$

KOSDAQ 자료의 GARCH모형에서는 제1주성분 C_1 은 ϵ_{t-3}^2 을 중심으로 대칭을 이루고 있다. 즉, t 시점에서 제1주성분 C_1 은 3시점 전의 수익률에 영향을 가장 많이 받고, 그 다음으로 2시점과 4시점 전의 수익률에 영향을 많이 받는다고 해석할 수 있다. 반대로 1시점과 5시점 떨어져 있을 때의 수익률은

Table 3.1. Covariance matrix between squared returns in KOSDAQ

	ϵ_{t-1}^2	ϵ_{t-2}^2	ϵ_{t-3}^2	ϵ_{t-4}^2	ϵ_{t-5}^2
ϵ_{t-1}^2	30.887955	15.987232	10.595063	7.292035	6.070518
ϵ_{t-2}^2	15.987232	30.893919	15.993188	10.600980	7.298000
ϵ_{t-3}^2	10.595063	15.993188	30.899846	15.998989	10.606940
ϵ_{t-4}^2	7.292035	10.600980	15.998989	30.905118	16.004922
ϵ_{t-5}^2	6.070518	7.298000	10.606940	16.004922	30.911084

Table 3.2. Covariance matrix between squared and signed returns in KOSDAQ

	ϵ_{t-1}^{+2}	ϵ_{t-2}^{+2}	ϵ_{t-3}^{+2}	ϵ_{t-4}^{+2}	ϵ_{t-5}^{+2}	ϵ_{t-1}^{-2}	ϵ_{t-2}^{-2}	ϵ_{t-3}^{-2}	ϵ_{t-4}^{-2}	ϵ_{t-5}^{-2}
ϵ_{t-1}^{+2}	26.14	8.79	3.70	1.16	0.86	0.001	0.28	0.63	2.16	1.44
ϵ_{t-2}^{+2}	8.79	26.15	8.80	3.70	1.16	5.84	0.001	0.28	0.63	2.16
ϵ_{t-3}^{+2}	3.70	8.80	26.15	8.80	3.70	5.40	5.84	0.002	0.28	0.63
ϵ_{t-4}^{+2}	1.16	3.70	8.80	26.15	8.80	2.88	5.40	5.84	0.003	0.29
ϵ_{t-5}^{+2}	0.86	1.16	3.70	8.80	26.15	3.17	2.88	5.40	5.84	0.004
ϵ_{t-1}^{-2}	0.001	5.84	5.40	2.88	3.17	4.45	0.98	0.85	1.11	0.62
ϵ_{t-2}^{-2}	0.28	0.001	5.84	5.40	2.88	0.98	4.45	0.98	0.85	1.11
ϵ_{t-3}^{-2}	0.63	0.28	0.002	5.84	5.40	0.85	0.98	4.45	0.98	0.85
ϵ_{t-4}^{-2}	2.16	0.63	0.28	0.003	5.84	1.11	0.85	0.98	4.45	0.98
ϵ_{t-5}^{-2}	1.44	2.16	0.63	0.29	0.004	0.62	1.11	0.85	0.98	4.45

C_1 에 가장 영향을 적게 받는다고 결론을 내릴 수 있다. 제1주성분 C_1 과 GARCH(1,1)모형에 의한 변동성 h_t 를 적합한 식은 다음과 같으며

$$h_t = 0.51806 + 0.3324 \left(\sqrt{C_1} \right)^2. \quad (3.1)$$

결정계수 값은 0.91이다. 같은 방법으로 T-GARCH모형에 대하여 적합 시켰다. T-GARCH모형의 PCA in NIC의 성분은 $\epsilon_{t-1}^{+2}, \epsilon_{t-2}^{+2}, \epsilon_{t-3}^{+2}, \epsilon_{t-4}^{+2}, \epsilon_{t-5}^{+2}, \epsilon_{t-1}^{-2}, \epsilon_{t-2}^{-2}, \epsilon_{t-3}^{-2}, \epsilon_{t-4}^{-2}, \epsilon_{t-5}^{-2}$, 총 10개이다. 이들이 표본 covariance matrix은 Table 3.2와 같다. 수익률의 분산은 정상성으로 인하여 시차에는 관계 없이 일정하다. 하지만, 수익률의 부호에 따라 그 값이 달라진다. 수익률이 양수인 경우에는 분산이 약 26.15로, 음수인 경우에는 4.45로 일정하다. 또한 부호만 같으면, 시차의 간격이 일정한 경우 그에 해당되는 공분산의 값이 일정하다. 이들의 covariance matrix를 이용하여 주성분 분석을 실시한 후, 그 결과를 보았다.

$$C_1 = 0.28781\epsilon_{t-1}^{+2} + 0.43476\epsilon_{t-2}^{+2} + 0.51583\epsilon_{t-3}^{+2} + 0.48928\epsilon_{t-4}^{+2} + 0.37587\epsilon_{t-5}^{+2} \\ + 0.1805\epsilon_{t-1}^{-2} + 0.15742\epsilon_{t-2}^{-2} + 0.12224\epsilon_{t-3}^{-2} + 0.08134\epsilon_{t-4}^{-2} + 0.04975\epsilon_{t-5}^{-2}.$$

KOSDAQ 자료에서 T-GARCH모형의 제1주성분을 보면, 수익률이 양수인 경우가 음수인 경우보다 변동성에 더 영향을 주는 것을 확인할 수 있다. 또한 수익률이 양수인 경우, $t-3$ 시차에서 가장 영향을 많이 받는다는 것을 알 수 있다. 이와 다르게 수익률이 음수인 경우에는 t 시점에서 가장 가까이 존재하는 $t-1$ 시점에서 가장 큰 영향을 주고, 차례대로 영향력이 감소하면서 마지막으로 $t-5$ 시점에서 변동성에 가장 적게 영향을 준다고 볼 수 있다. 제1주성분 C_1 과 T-GARCH(1,1)모형에 의한 변동성 h_t 를 적합한 식은 다음과 같으며

$$h_t = 0.65455 + 0.47942 \left(\sqrt{C_1} \right)^2, \quad (3.2)$$

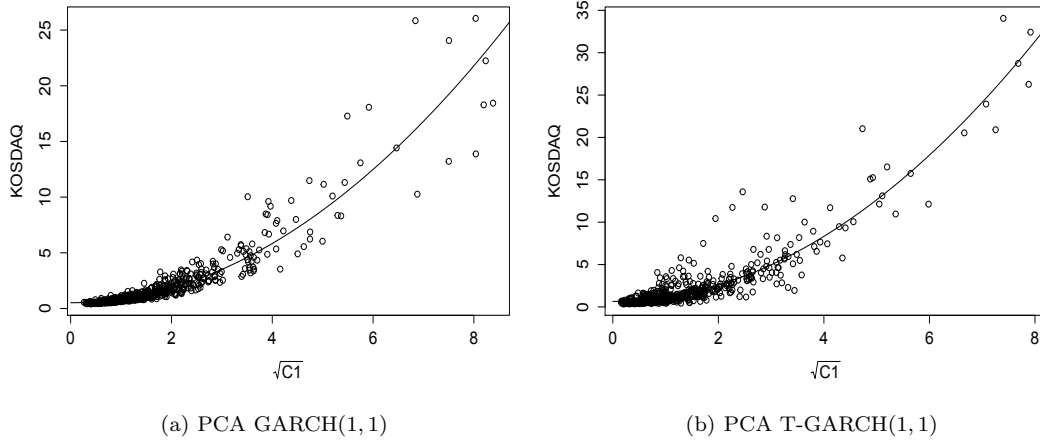


Figure 3.4. PCA in NIC of KOSDAQ

결정계수는 0.88이다. 식 (3.1)과 식 (3.2)를 토대로 PCA in NIC를 그려보면, Figure 3.4와 같다. GARCH모형의 PCA in NIC의 결과로 결정계수가 0.91로 상당히 높은 값이 나왔다. 이는 GARCH모형의 변동성 h_t 를 $t-1$ 시점부터 $t-5$ 시점까지 시차가 서로 다른 5가지의 수익률로 설명하기에 설명력이 강하다는 의미이다. PCA in NIC를 통해서 적은 양의 정보 손실로 5차원의 자료를 2차원으로 축소했음을 볼 수 있다. 이와 유사하게 T-GARCH모형에서의 PCA in NIC는 결정 계수 값이 0.88로 이 또한 상당히 적합이 잘 되었음을 볼 수 있다. 이는 10차원의 자료를 2차원으로 축약함으로써 적은 양의 정보 손실과 함께 좋은 결과를 얻을 수 있었다. 기존의 News Impact Curve에 비하여 PCA in NIC는 더 많은 시점의 수익률들을 포함하여 보기 쉽게 나타내었다. 또한 차원 축소라는 다변량 분석 기법을 사용하여 다차원에서 2차원으로 축소하여 보기 쉽게 나타내는 결과를 얻었다.

3.2. 사례분석 2 : 수익률 ϵ_t 이 조건부 t 인 경우의 KOSDAQ 자료

2010년 3월 2일부터 2013년 3월 2일까지의 749개의 KOSDAQ 데이터를 수익률이 t -분포를 따른다는 가정 하에 적합 시켜 보았다. 전체적인 분석절차는 조건부 정규분포와 유사하므로 주요 결과만 나열하기로 한다. 2D-NIC에 대한 추정식은 다음과 같다.

$$\begin{aligned}
 \text{GARCH ; } h_t &= 0.169464\epsilon_{t-1}^2 + 0.108107\epsilon_{t-2}^2 + 0.332751, \\
 \text{T-GARCH ; } h_t &= \begin{cases} 0.041817\epsilon_{t-1}^{+2} + 0.027140\epsilon_{t-2}^{+2} + 1.133135, & \text{if } \epsilon_{t-1} > 0 \text{ and } \epsilon_{t-2} > 0, \\ 0.041817\epsilon_{t-1}^{+2} + 0.161143\epsilon_{t-2}^{-2} + 1.133135, & \text{if } \epsilon_{t-1} > 0 \text{ and } \epsilon_{t-2} < 0, \\ 0.248284\epsilon_{t-1}^{-2} + 0.027140\epsilon_{t-2}^{+2} + 1.133135, & \text{if } \epsilon_{t-1} < 0 \text{ and } \epsilon_{t-2} > 0, \\ 0.248284\epsilon_{t-1}^{-2} + 0.161143\epsilon_{t-2}^{-2} + 1.133135, & \text{if } \epsilon_{t-1} < 0 \text{ and } \epsilon_{t-2} < 0. \end{cases}
 \end{aligned}$$

Figure 3.5(a)를 보면, GARCH모형은 $t-1$ 시점의 수익률과 $t-2$ 시점의 수익률에 대하여 변동성이 대칭으로 증가함을 볼 수 있다. T-GARCH모형의 경우에는 수익률들의 부호에 따라 비대칭적으로 증가하는데, 이들 중 $t-1$ 시점의 수익률과 $t-2$ 시점의 수익률 모두에서 음수의 부호일 때 가장 빠르게 증가된다. 이는 Figure 3.5(b)에서도 확인할 수 있다. GARCH(1,1)모형에서의 PCA in NIC를 보기 위하여, $\epsilon_{t-1}^2, \epsilon_{t-2}^2, \epsilon_{t-3}^2, \epsilon_{t-4}^2, \epsilon_{t-5}^2$ 간의 covariance matrix를 사용하여 주성분분석을 실시하였다. 제1주성

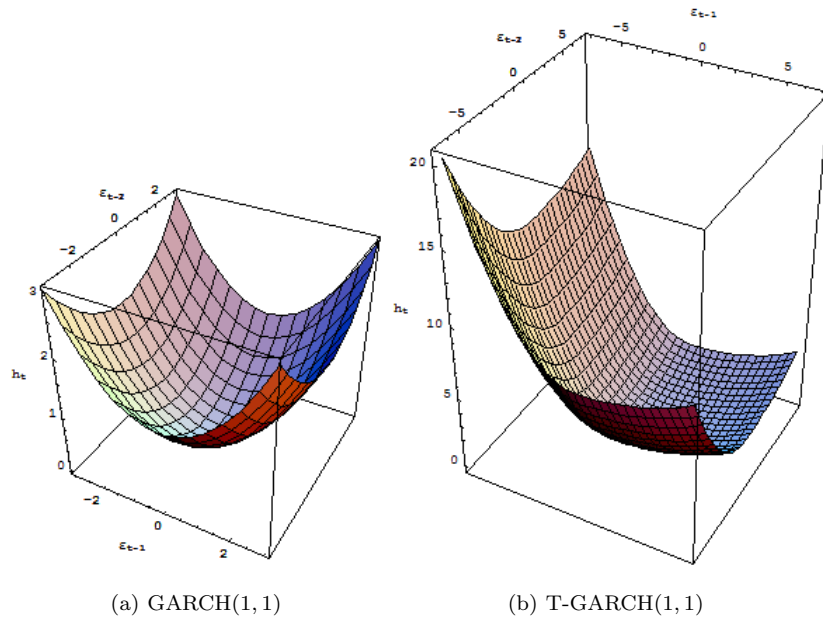


Figure 3.5. 2D-NIC of KOSDAQ (standardized t -distribution)

분 C_1 의 결과는 다음과 같다.

$$C_1 = 0.393549\epsilon_{t-1}^2 + 0.470159\epsilon_{t-2}^2 + 0.498113\epsilon_{t-3}^2 + 0.470167\epsilon_{t-4}^2 + 0.393569\epsilon_{t-5}^2.$$

KOSDAQ 자료의 GARCH모형에서는 제1주성분 C_1 는 계수들이 ϵ_{t-3}^2 을 중심으로 대칭을 이루고 있다. 이는 $t-3$ 시점 수익률이 C_1 에 가장 영향을 많이 받고, 그 다음으로 $t-2$ 시점과 $t-4$ 시점이 영향을 많이 준다고 해석할 수 있다. 반대로 다섯 시점 중 맨 끝인 $t-1$ 시점과 $t-5$ 시점의 수익률이 C_1 에 영향을 가장 적게 준다. 제1주성분 C_1 과 GARCH(1,1)모형에 의한 변동성 h_t 를 적합한 식은 다음과 같으며

$$h_t = 0.29769 + 0.19812 \left(\sqrt{C_1} \right)^2, \quad (3.3)$$

결정계수는 0.91로 분석되었다. 같은 방법으로 T-GARCH모형에 적용시킨 결과, 제1주성분 C_1 과 T-GARCH(1,1)모형의 변동성 h_t 사이의 식은 다음과 나타났으며

$$h_t = 0.33981 + 0.2957 \left(\sqrt{C_1} \right)^2, \quad (3.4)$$

결정계수는 0.88이다. GARCH모형의 PCA in NIC의 결정계수가 0.91로, T-GARCH모형의 결정계수는 0.88로 두 모형 모두 우수한 적합력을 갖는다. 이는 이전 시점의 수익률들이 변동성에 얼마나 영향을 주는지 PCA in NIC를 통해서 설명이 가능하다는 결과로 해석할 수 있다.

References

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

- Choi, M. S., Park, J. A. and Hwang, S. Y. (2012). Asymmetric GARCH processes featuring both threshold effect and bilinear structure, *Statistics & Probability Letters*, **82**, 419–426.
- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, **50**, 987–1008.
- Engle, R. F. and Ng, V. K. (1993). Measuring and testing the impact of news on volatility, *Journal of Finance*, **48**, 1749–1778.
- Rabemananjara, R. and Zakoian, J. M. (1993). Threshold ARCH models and asymmetries in volatility, *Journal of Applied Econometrics*, **8**, 31–49.
- Seong, W. H. (1997). *Applied Multivariate Analysis - Theory and Methods using SAS*, Tamjin, Seoul.

시계열 변동성 그래프의 개선

이정원^a · 윤재은^a · 황선영^{a,1}

^a숙명여자대학교 통계학과

(2013년 8월 12일 접수, 2013년 8월 26일 수정, 2013년 8월 26일 채택)

요약

News impact curves(NIC)는 1993년 Engle와 Ng에 의하여 제시되었으며, 이는 시계열 자료에서 발생하는 변동성을 시각적으로 나타내는데 용이하다. 본 논문에서는 기존의 NIC에서 더 나아가, 2차원 NIC(two dimensional NIC)와 주성분 NIC(PCA in NIC)를 제안하였으며, KOSDAQ 자료에서 적용하여 보았다.

주요용어: News Impact Curve, 주성분 분석, 변동성.

¹교신저자: (142-742) 서울특별시 용산구 청파로47길 100, 숙명여자대학교 통계학과, 교수.
E-mail: shwang@sookmyung.ac.kr