

# Multivariate Volatility Analysis via Canonical Correlations for Financial Time Series

Seung Yeon Lee<sup>a</sup> · S. Y. Hwang<sup>a,1</sup>

<sup>a</sup>Department of Statistics, Sookmyung Women's University

(Received November 16, 2014; Revised November 22, 2014; Accepted November 23, 2014)

---

## Abstract

Multivariate volatility is summarized through canonical correlation analysis (CCA). Along with the standard CCA, non-negative and sparse canonical correlation analysis (NSCCA) is introduced to make sure that volatility coefficients are non-negative and the number of coefficients in the volatility CCA is as small as possible. Various multivariate financial time series are analyzed to illustrate the main contribution of the paper.

Keywords: Canonical correlation analysis(CCA), multivariate volatility, non-negative and sparse canonical correlation analysis(NSCCA)

---

## 1. 서론

KOSDAQ, KOSPI, 환율과 같은 금융시계열의 시점  $t$ 에서의 변동성(volatility, 조건부 분산,  $h_t$ 로 표현)을 보면 수익률에 따라 값이 변하는 성질을 볼 수 있으며 과거 시점의 수익률에 따라 변동성이 달라진다고 알려져 있다. Engle과 Ng (1993)에 의해 처음 소개된 NIC(news impact curve)는  $t - 2$ 이전 시점까지의 정보는 상수로 고정시키고,  $t - 1$ 시점의 수익률과 그 다음 시점의 변동성인  $h_t$ 간의 관계를 그 그래프로 표현한 것이다. 이는  $t - 1$ 시점에서의 수익률이 한 시점 후의 변동성에 얼마나 영향을 주는지를 측정할 수 있게 한다(Yoon 등, 2014).

본 연구에서는 이변량 이상 다변량 금융시계열의 “변동성”을 연구하고자 한다. 이를 위해 정준상관분석 CCA(canonical correlation analysis)을 이용하여  $t - 1$ 시점의 “수익률”을  $x$ 축에,  $t$ 시점의 “변동성”을  $y$ 축에 나타내어 다변량으로 구성된 특정 산업의 변동성과 수익률에 대한 지표를 간단하게 NIC 그래프를 통해 나타냈다. 즉,  $t - 1$ 시점의 수익률의 선형결합인 정준상관분석 방법을 통하여 새로운 인공축인  $V_1$ (제1 정준상관변수)을 생성하고,  $t$ 시점의 변동성도 동일한 방법으로 또 다른 인공축인  $W_1$ (제1 정준상관변수)을 생성하였다.  $x$ 축에  $V_1$ 을,  $y$ 축에  $W_1$ 을 적용하여 나타낸 그래프를 CCA in NIC로 명명하기로 한다. 금융시계열자료에 대해 정준상관분석을 실시하면 수익률과 변동성이 음수로 나오는 경우가 발생한다. 이러한 경우 해석이 어렵다는 문제점이 존재한다. 이를 해결하기 위해 다양한 시점의 영향을 받는 수익률이 양의 값을 갖도록 제약조건을 주며 수익률에 크게 영향을 미치지 않는 시점을 0으로 간주하는 NSCCA(non-negative and sparse CCA)를 도입하였다. NSCCA에서는 일주일간의 거래 일을 고

---

<sup>1</sup>Corresponding author: Professor, Department of Statistics, Sookmyung Women's University, Yongsan-Gu, Seoul 142-742, Korea. E-mail: [shwang@sookmyung.ac.kr](mailto:shwang@sookmyung.ac.kr)

러한  $t - 5$ 시점까지의 수익률의 선형결합과 한 달간 거래일을 나타내는  $t - 22$ 시점까지 확장한 수익률의 선형결합을 통하여 인공축  $V_1$ (제1 정준상관변수)으로 나타낸 그래프인 NSCCA in NIC를 고려하였다. NSCCA 에서 인공축  $W_1$ 은 CCA in NIC와 마찬가지로  $t$ 시점의 변동성으로 이루어진 선형모형을 이용하였다. 다변량 기법인 정준상관분석은 종속변수군과 독립변수군 사이의 “상관구조”를 가장 잘 설명할 수 있는 선형방정식을 찾는 방법이다. 이 방법을 이용하여 다변량 금융시계열에 대한 차원 축소와 함께 특정 산업의 변동성과 수익률에 영향을 미치는 기업과 가중치에 대한 정보를 포함하는 지표를 만들 수 있을 것이며 이를 제시하기 위해 다양한 국내(다변량) 금융시계열 자료에 적용해 보았다.

## 2. GARCH 모형

금융시계열은 변동성 집중(volatility clustering), 비대칭성 그리고 급침(leptokurtic)분포를 따른다는 성질을 갖는다. 금융시계열에서 이와 같은 현상을 이해하기 위해 Engle (1982)은 ARCH 모형을 통해 금융시계열이 갖는 특징들을 모형화 하였으며, Bollerslev (1986)는 ARCH 모형을 일반화 시킨 GARCH 모형을 제안하였다. Bollerslev (1986)가 제안한 GARCH( $p, q$ ) 모형은 다음과 같다.

$$\begin{aligned} \epsilon_t &= \sqrt{h_t}e_t, \\ h_t &= \alpha_0 + \sum_{i=1}^p \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}, \quad (\alpha > 0, \alpha_i \geq 0, \beta_j \geq 0). \end{aligned}$$

여기서  $e_t$ 는 평균이 0이고 분산이 1인 i.i.d 확률변수이다. 변동성 분석에 가장 널리 사용되는 모형은 GARCH(1,1)모형이며 본 연구에서는 이 모형을 정준상관분석에 이용하였다.

### 2.1. News Impact Curve

Engle과 Ng (1993)가 제안한 NIC(news impact curve)는 변동성을 그래프로 표현하는 방법으로서 이를 통해 변동성의 함수형태 및 비대칭적 효과를 측정하는 도구로써 널리 이용되고 있다(Lee 등, 2013). NIC에서는  $t - 1$  시점의 수익률인  $\epsilon_{t-1}$ 를 제외한 모든 시점을 상수로 변환한 후,  $t$ 시점에서의 변동성  $h_t$ 와  $t - 1$  시점에서의 수익률간의 관계를 시각적으로 표현한 것이다. NIC에서 수익률  $\epsilon_{t-1}$ 은 가로축, 변동성  $h_t$ 는 세로축에 나타낸다. 비대칭-GARCH 모형의 경우, 좋은 뉴스는  $\epsilon_{t-1}^+$ 부분에서 변동성의 기울기를, 나쁜 뉴스는  $\epsilon_{t-1}^-$ 부분에서 변동성의 기울기를 통해 시각적으로 확인할 수 있다. 만약 레버리지(비대칭)효과가 존재 한다면  $\epsilon_{t-1}^-$ 의 기울기가  $\epsilon_{t-1}^+$ 의 기울기보다 더 커지게 된다. 본 논문에서는 정준상관분석을 통한 NIC 를 이용하기 위해 GARCH(1,1)모형만을 적합하였으며 이때의 NIC 방정식은 다음과 같다.

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

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

### 2.2. CCA in NIC

정준상관 분석은 여러 변수들을 두 변수 집단인 종속변수집단과 독립변수집단으로 분류한 후 두 변수 집단 사이의 상관구조를 가장 잘 설명할 수 있는 선형결합을 찾는 방법이다. 정준상관분석의 가장 큰 목표

는 서로 다른 집단들의 기본적인 관계를 발견하는 것이다. 같은 분야의 서로 다른 기업의 주가 데이터를 이용하여 특정 산업의 변동성과 수익률의 관계를 파악하고자 한다. 정준상관분석에 대한 내용은 성우현 (1997)의 제 6장을 참고하여 정리하였다.

두 개의 집단으로 구분되는  $p + q$ 개 확률변수들이 두 개의 확률벡터  $Y^T = [y_1, y_2, \dots, y_p]$ ,  $X^T = [x_1, x_2, \dots, x_q]$ 로 표현된다고 하자. 공분산 행렬들로 구성되는 행렬  $S_{yy}^{-1/2}S_{yx}S_{xx}^{-1}S_{xy}S_{yy}^{-1/2}$ 의  $p$ 개 고유값 중  $k$ 번째 고유값  $\hat{\lambda}_k$ 에 대응되는 고유벡터가  $\hat{e}_k$ 이고, 행렬  $S_{yy}^{-1/2}S_{yx}S_{xx}^{-1}S_{xy}S_{yy}^{-1/2}$ 의  $q$ 개 고유값 중  $k$ 번째 고유값  $\hat{\lambda}_k$ 에 대응되는 고유벡터가  $\hat{f}_k$ 일 때  $k$ 번째 표본 정준변수 짝은 다음과 같이 유도된다.

$$\begin{aligned}\hat{W}_k &= \hat{a}_k^T Y = \hat{e}_k^T S_{yy}^{-1/2} Y, \\ \hat{V}_k &= \hat{b}_k^T X = \hat{f}_k^T S_{xx}^{-1/2} X.\end{aligned}$$

$k$ 번째 정준표본상관계수는  $r(\hat{W}_k, \hat{V}_k) = \hat{\rho}_k^*$ 라고 할 때  $\hat{\rho}_k^* = \sqrt{\hat{\lambda}_k}$ 의 관계가 성립된다. 또한 표본정준변수는 단위 표본분산을 갖고 두 표본정준변수사이에는 서로 독립이므로 표본상관계수는 0이 된다. 변수가 표준화 되었다면 표본상관행렬을 다음과 같이 표현할 수 있다.

$$R_{uu} = \begin{bmatrix} R_{yy} & R_{yx} \\ R_{xy} & R_{xx} \end{bmatrix}$$

이때, 행렬  $R_{yy}^{-1/2}R_{yx}R_{xx}^{-1}R_{xy}R_{yy}^{-1/2}$ 와  $R_{xx}^{-1/2}R_{xy}R_{yy}^{-1}R_{yx}R_{xx}^{-1/2}$ 에서 구한 고유값  $\hat{\lambda}_k$ 에 대응되는 고유벡터  $\hat{e}_k$ 와  $\hat{f}_k$ 를 구하게 된다. 따라서 표준화 변수를 사용하였을 때  $k$ 번째 표본정준변수 짝은 다음과 같이 나타낼 수 있으며 본 논문에서는 아래의 식을 이용한다.

$$\begin{aligned}\hat{W}_k &= \hat{a}_k^T Y^* = \hat{e}_k^T p_{yy}^{-1/2} Y^*, \\ \hat{V}_k &= \hat{b}_k^T X^* = \hat{f}_k^T p_{xx}^{-1/2} X^*.\end{aligned}$$

여기서  $Y^*$ 와  $X^*$ 는 표준화된 벡터이다. 정준상관 NIC는 기존의 NIC와 똑같이 2차원 좌표축에 나타내지만, 가로축은  $t - 1$  시점의 특정 분야 기업들의 수익률들로 이루어진 제1 정준상관계수  $V_1$ 으로, 세로축은 특정 분야 기업들의  $t$ 시점에서의 변동성들로 이루어진 제1 정준상관계수  $W_1$ 을 이용한다는 점에서 차이가 있다. 우리가 본 연구에서 사용하게 될 제1 정준상관계수인  $V_1$ 과  $W_1$ 은 가장 상관성이 높은 선형결합으로 기존의 변동성 시계열 모형에서는 찾을 수 없었던 수익률과 변동성 사이의 연관성을 제공할 수 있을 것이다.

### 2.3. NSCCA in NIC

금융시계열데이터에 정준상관분석을 실시하면 수익률과 변동성에 대한 정준상관변수의 일부 계수가 음수로 나타나는 경우가 있어 해석상 어려움이 생길 수 있다. 이러한 해석상의 문제점을 해결하고, 변동성에 영향을 주는 최대한 많은 과거 시점의 수익률을 함께 고려하기 위하여 비음(non-negative)이면서 희박한(sparse) 정준상관분석(NSCCA)을 실시한다. NSCCA에서 사용한 알고리즘은 Sigg 등 (2007)을 이용하였다.

앞에서 언급한 기존 정준상관분석에서 갖는 문제점을 해결하기 위해서 NSCCA에서는 몇 개의 정준상관계수의 가중치를 0으로 만드는 제약 조건을 적용하였다.

$$(\hat{a}, \hat{b}) = \operatorname{argmax}_{a,b} \operatorname{corr}(Xa, Yb). \quad (2.1)$$

비음의 정준상관 가중치를 만들기 위해 식 (2.1)에서  $a_i \geq 0, b_i \geq 0$ 라는 제약을 둔다. 아래의 식 (2.2)은 최소제곱법을 이용하기 위해 식 (2.1)을 변형시킨 것이며, 이를 반복 적용하여 최적의  $\hat{a}, \hat{b}$ 를 찾는다.

$$(\hat{a}, \hat{b}) = \operatorname{argmin} \| Xa - Yb \|^2. \quad (2.2)$$

$\hat{a}$ 이 주어진 상황에서, 다음의 식 (2.3)을 이용하여 최소제곱평균오차기준(minimum mean-squared error criterion)을 이용하여 회귀계수  $b$ 를 찾을 수 있다.

$$\hat{b} = \operatorname{argmin}_b \| X\hat{a} - Yb \|^2. \quad (2.3)$$

식 (2.3)을 이용하여 식 (2.4)와 같은 방법으로  $\hat{b}$ 를 재추정해 나간다.

$$\hat{b} \leftarrow \frac{\hat{b}}{\| Y\hat{b} \|}. \quad (2.4)$$

$\hat{a}^{(t)}$ 가 주어진 상태에서  $(t+1)$ 번 반복 되었을 경우, 위 식 (2.3)과 (2.4)를 이용하여  $\hat{a}^{(t)}$ 에 대응하는  $\hat{b}^{(t+1)}$ 를 찾을 수 있으며  $\hat{b}^{(t+1)}$ 가 결정되면, 이에 최적화 된  $\hat{a}^{(t+1)}$ 를 찾을 수 있다.

알고리즘은 다음과 같다.

step 1.  $r = X\hat{a}, b = 0$  으로 시작한다.

step 2.  $Y$ 행렬로부터  $r$ 과 가장 큰 양의 상관관계를 갖는  $j$ 번째 열의  $y_j$  벡터를 찾는다.

step 3.  $b_j \leftarrow b_j + \epsilon$  으로 대체한다.

step 4.  $r = r - \epsilon b_j$ 로 대체한 후 step 2와 step 3를 사전에 설정한 노름(norm)의 조건을 위배하거나 양의 상관관계를 갖는 열이 없을 때 까지 계속 실시한다.

위 알고리즘을 통해

$$\hat{b}_{(k)} = \operatorname{argmin}_b \| X\hat{a} - Yb \|^2 + \lambda b^T O b,$$

$$\text{s.t. } \sum_{g=1}^G \| b_{[g]} \| \leq t \wedge b_j \geq 0 \quad \forall j, \quad \text{where } O = \sum_{l < k} R_{yy} \hat{b}_{(l)} \hat{b}_{(l)}^T R_{yy}$$

인 NSCCA를 찾을 수 있다. 자세한 내용은 Sigg 등 (2007)을 참조하기 바란다.

### 3. 국내금융시계열에의 응용

본 절에서는 다변량 시계열 중에서 특히 이변량 시계열에 초점을 맞추고자 한다. 2009년 1월 2일부터 2014년 3월 31일까지의 은행산업(국민, 신한), 건설업(대림, GS), 자동차산업(KIA, 현대), 전자산업(LG, 삼성)의 대표적인 두 개의 기업들과 KOSDAQ, KOSPI의 주가 데이터를 이용하여 CCA, NSCCA에 적용 시킨 후 NIC를 이용하여 특정 산업을 대표하는 지표를 시각적으로 나타내었다. 이를 통해 각 기업의 변동성과 수익률을 따로 살펴보는 것이 아니라 특정 산업의 변동성과 수익률에 영향을 주는 기업과 시점을 아는 것과 동시에 지표를 만들어 한눈에 볼 수 있다는 장점이 있다. 사용한 프로그램은 SAS와 R의 nscancor 패키지이다. 각 주가 데이터를 백분위 로그수익률( $R_t =$

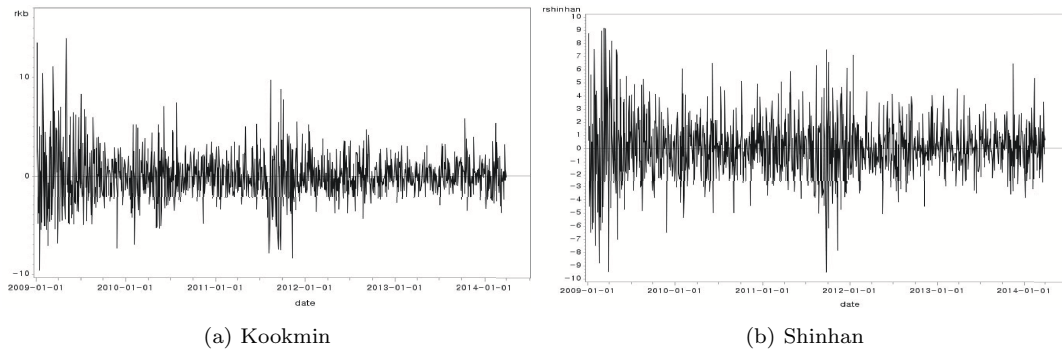


Figure 3.1. Time series plot for bank sector

Table 3.1. Estimated ARMA-GARCH(1,1)

산업분야	기업명	적합결과
은행산업	국민은행	AR(2)-GARCH(1,1) $h_t=0.0372+0.0546h_{t-1}+0.9382\epsilon_{t-1}^2$
	신한은행	AR(1)-GARCH(1,1) $h_t=0.0512+0.0444h_{t-1}+0.9432\epsilon_{t-1}^2$
건설산업	대림건설	AR(1)-GARCH(1,1) $h_t=0.3636+0.0790h_{t-1}+0.8785\epsilon_{t-1}^2$
	GS건설	AR(1)-GARCH(1,1) $h_t=0.4722+0.0661h_{t-1}+0.8788\epsilon_{t-1}^2$
자동차산업	기아자동차	ARMA(1,1)-GARCH(1,1) $h_t=0.1633+0.6618h_{t-1}-0.6618\epsilon_{t-1}^2$
	현대자동차	ARMA(1,1)-GARCH(1,1) $h_t=0.1790+0.0577h_{t-1}+0.9090\epsilon_{t-1}^2$
전자산업	LG전자	AR(1)-GARCH(1,1) $h_t=0.0147+0.0333h_{t-1}+0.9635\epsilon_{t-1}^2$
	삼성전자	ARMA(1,1)-GARCH(1,1) $h_t=0.0782+0.0419h_{t-1}+0.9357\epsilon_{t-1}^2$

$\ln(P_t/P_{t-1} \times 100)$ )로 변형 한 후 분석에 이용하였다.  $R_t$ 로 변형된 시도표 중 은행 산업의 시도표를 보면 Figure 3.1과 같다.

신한은행과 국민은행의 시도표를 살펴보면 변동의 폭은 다르지만 비슷한 형태를 보이는 것을 알 수 있다. 나머지 산업들 역시 시도표가 비슷한 형태를 보이므로 생략하기로 한다.

### 3.1. 사례분석 1: CCA in NIC

2009년 1월 2일부터 2014년 3월 31일까지의 1307개의 총 8개 수익률 데이터를 ARMA-GARCH(1,1)모형에 적합하였다. 추정 식은 Table 3.1과 같다. 추정된 ARMA-GARCH(1,1) 모형을 이용하여 각 산업에 대해 정준상관분석을 실시한 결과인 Table 3.2를 살펴보자. 각 기업은 약어(국민은행(kb), 신한

Table 3.2. CCA-NIC

산업분야	CCA 적합결과	NIC 적합결과
은행산업	$\hat{W}_1=0.6909h_{t, kb}^*+0.3267h_{t, shinhan}^*$ $\hat{V}_1=0.4509\epsilon_{t-1, kb}^*+0.6319\epsilon_{t-1, shinhan}^*$	$w_1 = 0.4909v_1^2$
건설산업	$\hat{W}_1=0.9704h_{t, dl}^*+0.0414h_{t, gs}^*$ $\hat{V}_1=0.9034\epsilon_{t-1, dl}^*+0.1531\epsilon_{t-1, gs}^*$	$w_1 = 0.5298v_1^2$
자동차산업	$\hat{W}_1=0.2909h_{t, hcar}^*+0.7574h_{t, kia}^*$ $\hat{V}_1=0.0778\epsilon_{t-1, hcar}^*+0.9475\epsilon_{t-1, kia}^*$	$w_1 = 0.4272v_1^2$
전자산업	$\hat{W}_1=0.0434h_{t, lg}^*+0.9700h_{t, samsung}^*$ $\hat{V}_1=0.3579\epsilon_{t-1, lg}^*+0.8489\epsilon_{t-1, samsung}^*$	$w_1 = 0.4320v_1^2$

은행(shinhan), 대림건설(dl), gs건설(gs), 현대자동차(hcar), 기아자동차(kia), LG 전자(lg), 삼성전자(samsung))로 나타냈으며, CCA 적합결과에서  $h_t$ 와  $\epsilon_{t-1}$ 는 분산이 1이 되도록 표준화 시킨 표준 정준상관으로  $h_t^*$ 와  $\epsilon_{t-1}^*$ 로 표현하였다. 은행 산업의 정준상관결과를 먼저 살펴보면, 국민은행의 수익률(0.4509)과 신한은행의 수익률(0.6319) 중 신한은행의 수익률에 더 큰 가중치를 두고 있으며, 변동성에 있어서는 국민은행이 신한은행에 비해 큰 가중치를 두고 있다. 은행 산업의 경우 변동성과 수익률간의 가중치를 주는 기업이 다른 것을 볼 수 있다. 전자 산업의 정준상관결과에서 변동성을 먼저 살펴보면, 삼성전자(0.97)가 LG전자(0.0434)에 비해 전자 산업에 있어서 20배에 가까운 매우 큰 영향을 미치고 있는 것으로 보이며, 수익률에 있어서도 삼성전자는 0.8489, LG전자는 0.3579로 삼성전자가 LG전자에 비해 2배 이상의 가중치를 두는 것을 알 수 있다. 변동성과 수익률에 있어서 영향을 미치는 기업이 달랐던 은행산업과는 달리 전자산업에서는 특정 기업이 변동성과 수익률에 매우 큰 영향을 미치고 있는 것으로 나타났다. 건설 산업과 자동차산업의 정준상관결과 또한 전자산업과 같이 변동성과 수익률 모두에서 특정 기업이 큰 가중치를 차지하는 것을 볼 수 있다.

Table 3.2에서 NIC 적합결과를 살펴보자. NIC 모형적합결과 상수는 유의하지 않아 상수를 제거한 모형을 최종 모형으로 결정하여 NIC에 적합하였다. 각 산업에 대해 NIC를 적합시킨 그래프는 Figure 3.2와 같다. 네 산업 모두 정준상관계수가 양수로 나왔으며 대칭인 GARCH(1,1) 모형을 이용하여 정준상관분석을 실시하였기 때문에 0을 기준으로 대칭인 2차 형식을 띄고 있는 것을 알 수 있다. 이를 통해 정준상관분석을 이용해 구한 수익률의 선형결합이 변동성의 선형결합에 영향을 준다는 결론을 얻을 수 있다.

### 3.2. 사례분석 2 : NSCCA in NIC

2009년 1월 2일부터 2013년 3월 2일까지의 1288개의 KOSDAQ과 KOSPI 자료를 일별 수익률  $R_t$ 로 변형 시킨 후 분석하였다. KOSPI와 KOSDAQ 수익률의 ARMA-GARCH(1,1) 적합 결과는 다음과 같다. 정준상관분석 시 수익률의 계수가 음수가 되면 해석하기 어렵다는 문제점이 발생한다. 또한 수익률은  $t-1$  시점 뿐 만 아니라  $t-1, t-2, \dots$  등 무수히 많은 시점에 영향을 받는다. 본 예제에서는 KOSPI와 KOSDAQ의  $t-5$ (일주일)시점과  $t-22$ (한달)시점까지의 데이터를 이용하여 수익률이 음수가 되지 않으며 크게 영향을 미치지 않는 변수들을 영으로 만드는 비음이면서 희박한 정준상관분석, 즉 NSCCA을 통해 NIC를 시각화 하려 한다.

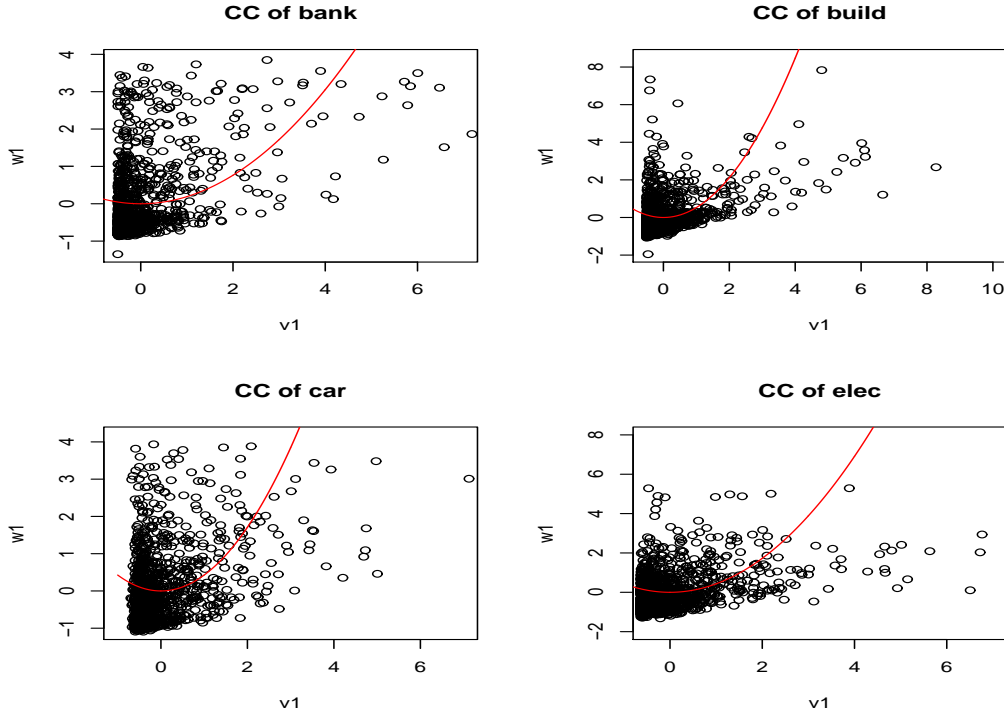


Figure 3.2. NIC

Table 3.3. ARMA-GARCH(1,1) for KOSPI and KOSDAQ

적합결과	
KOSPI	ARMA(1,1)-GARCH(1,1) $h_t=0.0158+0.0694h_{t-1}+0.9198\epsilon_{t-1}^2$
KOSDAQ	AR(1)-GARCH(1,1) $h_t=0.1258+0.2577h_{t-1}+0.6903\epsilon_{t-1}^2$

3.3.  $t - 5$  시점까지의 KOSPI와 KOSDAQ 자료

위에서 적합시킨 ARMA-GARCH(1,1) 모형의 결과를 이용하여 KOSPI와 KOSDAQ에서 총  $t - 5$  시점까지 총 10개의 수익률에서 4개의 수익률을 성분으로 하는 NSCCA를 실시하여 제1 정준상관계수를 추출하였다. 추출된 제1 정준상관계수를  $W_1, V_1$ 이라 하며 결과는 다음과 같다.

$$\hat{W}_1 = 0.0262h_{t,kospi}^* + 0.0023h_{t,kosdaq}^*,$$

$$\hat{V}_1 = 0.00001706\epsilon_{t-2,kosdaq}^* + 0.00003642\epsilon_{t-5,kosdaq}^* + 0.022642\epsilon_{t-1,kospi}^* + 0.00006826\epsilon_{t-3,kospi}^*.$$

이를 통해 KOSPI가 KOSDAQ에 비해 변동성에서 큰 영향을 미치는 것을 확인 할 수 있었으며 이는 주가 시장의 규모가 크며 KOSDAQ에 비해 상대적으로 많은 투자를 받기 때문으로 보인다. 수익률로 이루어진 제1 정준상관  $\hat{V}_1$ 을 보면  $\epsilon_{t-1,kospi}$ 의 수익률에 영향을 가장 많이 받고, 그 다음으로는

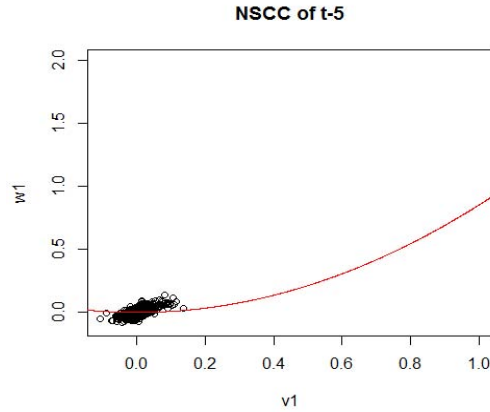


Figure 3.3. NSCCA ARMA-GARCH(1,1) upto  $t - 5$

$\epsilon_{t-3, \text{kospi}}$ 에 영향을 많이 받으며 상대적으로 KOSDAQ의 수익률은 KOSPI의 수익률에 비해 영향을 많이 받지 않는 것을 알 수 있다.

NSCCA에 의한 정준상관분석 적합 식은 식 (3.1)과 같으며 모형적합결과 상수는 유의하지 않아 상수를 제거한 모형을 최종 모형으로 결정하였다. 이때의 결정계수는 0.48 이다.

$$w_1 = 0.85306v_1^2 \quad (3.1)$$

식 (3.1)를 토대로 NSCCA in NIC를 그려보면 Figure 3.3과 같다.

### 3.4. $t - 22$ 시점까지의 KOSPI와 KOSDAQ 자료

또한  $t - 1$  시점부터  $t - 22$  시점까지의 수익률들(한 달, 총 22개)에 대해 NSCCA 분석을 실시 한 후 NIC를 그려보았다. 총 44 개의 수익률에서 11개의 수익률들을 성분으로 하는 NSCCA를 실시하여 제1 정준상관계수를 추출하였다. 추출된 제1 정준상관을  $W_1, V_1$ 이라 하고 결과는 다음과 같다.

$$\begin{aligned} \hat{W}_1 &= 0.02246h_{t, \text{kospi}}^* + 0.007213h_{t, \text{kosdaq}}^* \\ \hat{V}_1 &= 0.004277\epsilon_{t-1, \text{kosdaq}}^* + 0.000263\epsilon_{t-2, \text{kosdaq}}^* + 0.0003642\epsilon_{t-5, \text{kosdaq}}^* + 0.0000705\epsilon_{t-11, \text{kosdaq}}^* \\ &\quad + 0.0001244\epsilon_{t-18, \text{kosdaq}}^* + 0.01901\epsilon_{t-1, \text{kospi}}^* + 0.0004334\epsilon_{t-3, \text{kospi}}^* + 0.0002601\epsilon_{t-4, \text{kospi}}^* \\ &\quad + 0.00001512\epsilon_{t-10, \text{kospi}}^* + 0.0001479\epsilon_{t-17, \text{kospi}}^* + 0.0001096\epsilon_{t-18, \text{kospi}}^* \end{aligned}$$

$t - 5$  시점과 마찬가지로 변동성에 있어서 KOSPI가 KOSDAQ에 비해 더 큰 영향력을 미치는 것을 알 수 있다. 수익률을 살펴보면, 총 44개의 수익률 중 11개의 수익률만  $t$ 시점의 변동성에 영향을 미치는 것을 알 수 있다. 추출된 11개의 수익률 중 KOSPI에서는  $t - 1, t - 3, t - 4, t - 10, t - 17, t - 18$ 로 총 6개의 시점이 영향을 미친다는 것을 알 수 있었으며, KOSDAQ에서는  $t - 1, t - 2, t - 5, t - 11, t - 18$ 로 총 5개의 시점이 영향을 미친다는 것을 알 수 있었다. 또한 KOSPI와 KOSDAQ 모두  $t - 1$ 시점과  $t - 18$ 시점의 수익률에 영향을 받았으며, KOSPI와 KOSDAQ 모두  $t - 1$ 시점에서 가장 큰 영향을 받은 것을 확인 할 수 있었다.  $t - 22$  시점까지의 수익률을 이용한 NSCCA 정준상관분석 적합 식은 식



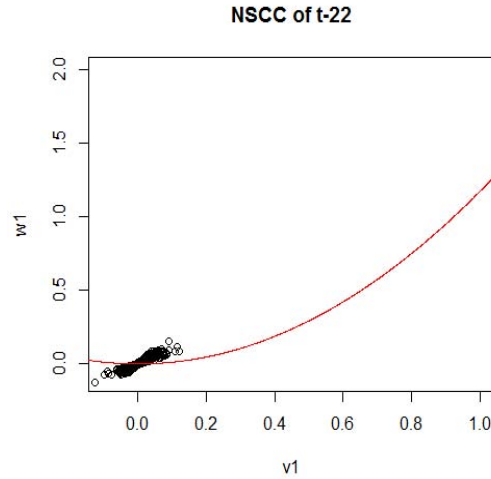


Figure 3.4. NSCCA ARMA-GARCH(1,1) upto  $t - 22$

(3.2)와 같다.

$$w_1 = 0.000688 + 1.17372v_1^2 \quad (3.2)$$

식 (3.2)를 토대로 NSCCA in NIC를 그려보면 Figure 3.4와 같다.

#### 4. 결론

본 연구에서는 다변량 금융시계열 자료를 이용하여 각 모형에 맞는 ARMA-GARCH모형을 적합한 후 정준상관분석을 이용하여 각 산업별로 변동성과 수익률과의 관계를 NIC를 통해 시각화 하였다. 요인분석(factor analysis)과 주성분분석(principal component analysis)을 이용한 차원축소는 이미 많은 연구가 보고 되고 있다. PCA를 이용하여 차원을 축소한 결과를  $x$ 축으로,  $y$ 축은  $h_t$  시점만의 변동성을 이용하여 NIC를 그린 Lee 등 (2013)과 달리 본 연구에서는  $x$ 축과  $y$ 축 모두 정준상관분석을 이용하여 차원 축소를 하였다. 이러한 시도를 통해 같은 산업에 속한 기업들의 변동성과 수익률을 나타내는 지표를 통해 특정 산업에서 영향을 미치는 기업을 확인 해 볼 수 있었다. 또한 NSCCA를 이용하여  $t - 1$  시점 뿐 아니라  $t - 5$  시점과  $t - 22$  시점까지 즉, 일주일 혹은 한 달 동안 어느 시점의 수익률이 변동성에 영향을 주었는지 알아 볼 수 있었다.

실증분석의 자료는 2009년 1월 2일부터 2014년 3월 31일까지의 로그수익률을 사용하였으며 수익률과 변동성에 대한 지표 생성과 해석의 용의성을 위해 양의 정준상관계수가 나오는 것에 초점을 맞춰 연구하였다.  $t - 1$  시점의 수익률만을 가지고 분석한 CCA와 다양한 시점(일주일과 한 달 간)의 수익률에 대해 실시한 NSCCA를 확인한 결과,  $t - 1$  시점의 수익률이 가장 큰 가중치를 가지는 것을 확인 할 수 있었다. 또한 변동성의 가중치는 기업의 규모에 의해 영향을 받는다는 것을 확인 할 수 있었다.

본 연구에서는 이변량 금융 시계열 자료를 통해서만 예시를 하였지만, 이변량 이상의 자료에 대해서도 동일한 분석을 실시 할 수 있으며, 본 논문에서 이용한 GARCH(1,1) 모형 뿐 아니라 현재 연구되고 있는 비대칭(asymmetric)-GARCH 모형에 대해서도 확장 연구를 할 수 있을 것으로 기대한다.

## References

- 성웅현 (1997). 응용 다변량분석-이론, 방법론, SAS활용, 탐진
- 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.
- Hwang, S. Y., Yoon, J. E. and Lee, J. W. (2013). A graphical improvement in volatility analysis for financial series, *Korean Journal of Applied Statistics*, **26**, 785–796.
- Buhmann, J., Fischer, B., Roth, V. and Sigg, C. (2007). Nonnegative CCA for audiovisual Source Separation, *Machine Learning for Signal Processing*, 253–258.
- Hwang, S. Y., Lee, J. W. and Yoon, J. E. (2014). News impact curve of volatility for asymmetric GARCH via LASSO, *Korean Journal of Applied Statistics*, **27**, 159–168.

# 정준상관분석을 통한 다변량 금융시계열의 변동성 분석

이승연<sup>a</sup> · 황선영<sup>a,1</sup>

<sup>a</sup>숙명여자대학교 통계학과

(2014년 11월 16일 접수, 2014년 11월 22일 수정, 2014년 11월 23일 채택)

---

## 요약

다변량 금융시계열의 변동성분석을 다변량 기법인 정준상관분석(canonocal correaltion analysis) 을 이용해 분석하였다. 변동성의 특성상 계수들이 비음(non-negative)인 정준상관분석, 즉, non-negative and sparse canonical correlation analysis (NSCCA)를 이용해 보았다. 본 논문은 다변량 시계열의 변동성 커브에 대해 연구하고 있으며 제시된 방법론을 이변량 주성분분석을 통해 예시해 보았다.

주요용어: 정준상관분석, 다변량 변동성, 비음(non-negative) 정준상관분석.

---

<sup>1</sup>교신저자: (142-742) 서울특별시 용산구 청파로 47길 100, 숙명여자대학교 통계학과, 교수.  
E-mail: shwang@sookmyung.ac.kr