

비선형 평균 일반화 이분산 자기회귀모형의 추정[†]

심주용¹ · 이장택²

¹대구가톨릭대학교 응용통계학과

²단국대학교 정보통계학과

접수 2010년 5월 13일, 수정 2010년 8월 12일, 게재확정 2010년 8월 28일

요약

최소제곱 서포트벡터기계는 비선형회귀분석과 분류에 널리 쓰이는 커널기법이다. 본 논문에서는 금융시계열자료의 평균 및 변동성을 추정하기 위하여 평균의 추정 방법으로는 가중최소제곱 서포트벡터기계, 변동성의 추정 방법으로는 최소제곱 서포트벡터기계를 사용하는 비선형 평균 일반화 이분산 자기회귀모형을 제안한다. 제안된 모형은 선형 일반화 이분산 자기회귀모형 및 선형 평균 일반화 이분산 자기회귀모형보다 더 나은 추정 능력을 가진다는 것을 실제자료의 추정을 통하여 보였다.

주요어: 가중최소제곱 서포트벡터기계, 일반화 교차타당성함수, 일반화 이분산 자기회귀모형, 최소제곱 서포트벡터기계, 평균 일반화 이분산 자기회귀모형.

1. 서론

금융시계열의 평균 및 변동성에 대한 정확한 추정과 예측은 금융모형분석 또는 이론검정에 있어서 중요한 뿐만 아니라 실무에서 선물, 옵션 등의 파생상품을 거래하는 투자자에게도 매우 중요하다. 변동성군집 또는 두꺼운 꼬리의 분포를 갖는 시계열을 조건부 분산의 관점에서 Engle (1982)은 변동성 (volatility)을 갖는 금융시계열을 조건부 분산의 관점에서 모형화하기 위하여 ARCH (autoregressive conditional heteroscedasticity)모형을 제안하였다. ARCH모형은 자산수익률의 변동성의 특징을 표현하기 위해서 GARCH, EGARCH, IGARCH, TGARCH모형 등과 같은 조건부 분산모형으로 확대되었다. 이런 GARCH모형들은 실제 금융시계열자료를 분석하기위해 널리 사용되고 있는 데, GARCH모형 중 가장 대표적인 모형은 GARCH(1,1)모형이며 자세한 내용은 Bollerslev (1986), Najand와 Yung (1991)에 설명되어있다.

한편 Perez-Cruz 등 (2003)은 오차항을 정규분포로 가정하는 GARCH(1,1)모형은 꼬리가 두꺼운 분포를 따르는 수익률 시계열자료를 충분히 설명할 수 없는 경우가 많다고 하여 오차항이 정규분포를 따르지 않는 GARCH(1,1)모형의 추정법으로 SVM (support vector machine, Vapnik, 1998) 기법을 제안하였다. Hwang과 Shin (2010)은 자료의 평균이 상수임을 가정하여 비선형 GARCH모형 추정을 위해 커널기법을 적용하는 방법을 제안하였다. Domowitz와 Hakkio (1985)는 평균과 변동성의 변화를 동시에 추정하기 위한 GARCH-M (GARCH in the mean)모형을 제안하였다.

[†] 본 과제는 정보통신산업진흥원의 SW공학 요소기술 연구개발사업의 결과물임을 밝힙니다.

¹ (702-701) 경북 경산시 하양읍, 대구가톨릭대학교 응용통계학과, 겸임교수.

² 교신저자: (448-701) 경기도 용인시 수지구 죽전동 126번지, 단국대학교 정보통계학과, 교수.
E-mail: jtlee@dankook.ac.kr

본 논문에서는 자료의 평균이 상수가 아닌 경우, 비선형 GARCH-M모형을 추정하기 위한 Suykens와 Vanderwalle (1999)의 가중 및 비가중최소제곱 서포트벡터기계 (least squares support vector machine, LS-SVM)를 적용한 반복 알고리즘을 제안하고 기존의 선형 GARCH모형 및 선형 GARCH-M모형과 성능을 비교하고자 한다. LS-SVM 및 커널기법에 대한 자세한 내용은 Shim과 Hwang (2003), Hwang (2007) 그리고 Shim과 Lee (2009) 등에 설명되어 있다.

본 논문은 다음과 같이 구성되어 있다. 2절에서는 비선형 GARCH-M모형의 추정을 위하여 가중 및 비가중최소제곱 서포트벡터기계를 적용한 반복 알고리즘을 제안하고, 3절에서는 모형선택을 위한 일반화 교차타당성기법에 대해 설명한다. 4절에서는 실제자료인 S&P500지수 자료와 미국재무성채권 자료에 제안된 방법을 적용하여 그 결과를 설명하고 결론을 기술한다.

2. 비선형 GARCH(1,1)-M모형의 추정

만약 시계열자료 $\{y_t\}_{t=1}^n$ 가 선형 GARCH(1,1)-M모형을 따른다면 평균방정식과 분산방정식은 일반적으로 다음과 같이 정의된다.

$$\begin{aligned} \text{평균방정식 : } y_t &= \gamma_0 + \gamma_1 y_{t-1} + \gamma_2 \sigma_t + \varepsilon_t, \quad t = 1, \dots, n, \\ \text{분산방정식 : } \sigma_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned} \quad (2.1)$$

여기서 ε_t 는 일반적으로 $N(0, \sigma_t^2)$ 을 따른다고 가정한다. 일반적으로 GARCH모형의 평균방정식은 $y_t = \mu + \varepsilon_t$ 로 주어지며, 분산방정식의 모수 ω, α, β 는 조건부 분산이 양수이어야 하므로 $\omega > 0, \alpha, \beta \geq 0$ 을 만족해야 한다. 또한 GARCH(1,1)-M모형에서 분산이 양수이고 정상성 조건을 만족하기 위한 조건은 $\alpha + \beta < 1$ 이다.

주어진 시계열자료 $\{y_t\}_{t=1}^n$ 의 평균이 상수가 아닌 경우에 대해 다음과 같은 비선형 GARCH(1,1)-M모형을 고려한다.

$$\begin{aligned} \text{평균방정식 : } y_t &= \mu_t + \varepsilon_t = f_1(y_{t-1}, \sigma_t) + \varepsilon_t, \quad t = 1, \dots, n, \\ \text{분산방정식 : } \sigma_t^2 &= f_2(\varepsilon_{t-1}^2, \sigma_{t-1}^2) \end{aligned} \quad (2.2)$$

여기서 f_1 과 f_2 는 각각 $\mathbf{x}_{1t} = (y_{t-1}, \sigma_t)$ 과 $\mathbf{x}_{2t} = (\varepsilon_{t-1}^2, \sigma_{t-1}^2)$ 의 비선형 함수이고, ε_t 는 분산이 σ_t^2 인 0을 중심으로 대칭인 분포를 따르는 오차항이다.

2.1. 평균의 추정

평균방정식에서 σ_t^2 가 주어진 경우, μ_t 의 추정을 위해서 우리는 커널기법을 사용하여 평균함수 $\mu_t = \mu(\mathbf{x}_{1t})$ 를 추정하고자 한다. 커널기법은 μ_t 를 고차원 특징공간에서의 선형모형 $\mu_t = \boldsymbol{\omega}'_1 \boldsymbol{\phi}(\mathbf{x}_{1t}) + b_1$ 를 사용하여 추정한다. 이때 비선형 특징함수 $\boldsymbol{\phi}(\cdot) : R^d \rightarrow R^{df}$ 는 입력공간을 더 높은 차원의 특징공간으로 사상시킨다. 이제 $\boldsymbol{\omega}_1$ 과 b_1 의 추정값을 구하기 위하여 다음과 같이 가중 LS-SVM (weighted LS-SVM, Suykens 등, 2002)의 최적화 문제를 고려한다.

$$\begin{aligned} \min \frac{1}{2} \|\boldsymbol{\omega}_1\|^2 + \frac{C_1}{2} \sum_{t=1}^n \frac{1}{\sigma_t^2} \varepsilon_t^2 \\ \text{subject to } \varepsilon_t = y_t - \boldsymbol{\omega}'_1 \boldsymbol{\phi}(\mathbf{x}_{1t}) - b_1 \end{aligned} \quad (2.3)$$

여기서 $C_1 > 0$ 은 자료의 적합 (fitting)과 함수의 평탄성 (flatness)을 조절하는 벌칙상수이다. 식 (2.3)을 사용하면 라그랑제 함수 (Lagrangian function)는 다음 식 (2.4)와 같이 만들어지며, 이 경우 α_{1t} 는 라그랑제 배수 (Lagrange multipliers)이다.

$$L(\boldsymbol{\omega}_1, b_1, \boldsymbol{\varepsilon} : \boldsymbol{\alpha}_1) = \frac{1}{2} \|\boldsymbol{\omega}_1\|^2 + \frac{C_1}{2} \sum_{t=1}^n \frac{1}{\sigma_t^2} \varepsilon_t^2 - \sum_{t=1}^n \alpha_{1t} (\varepsilon_t - y_t + \boldsymbol{\omega}'_1 \boldsymbol{\phi}(\mathbf{x}_{1t}) + b_1) \quad (2.4)$$

KKT 조건 (Karush-Kuhn-Tucker conditions, Kuhn과 Tucker, 1951)에 의하여 식 (2.4)의 라그랑제 함수를 최소화하는 관련 모수들의 값은 다음 식 (2.5)와 같은 연립방정식으로부터 구해진다.

$$\begin{aligned} \frac{\partial L}{\partial \boldsymbol{\omega}_1} = \mathbf{0} &\rightarrow \boldsymbol{\omega}_1 = \sum_{t=1}^n \alpha_{1t} \boldsymbol{\phi}(\mathbf{x}_{1t}) \\ \frac{\partial L}{\partial b_1} = 0 &\rightarrow \sum_{t=1}^n \alpha_{1t} = 0 \\ \frac{\partial L}{\partial \varepsilon_t} = 0 &\rightarrow \varepsilon_t = \frac{\sigma_t^2}{C_1} \alpha_{1t}, t = 1, 2, \dots, n \\ \frac{\partial L}{\partial \alpha_{1t}} = 0 &\rightarrow \varepsilon_t - y_t + \boldsymbol{\omega}'_1 \boldsymbol{\phi}(\mathbf{x}_{1t}) + b_1 = 0, t = 1, 2, \dots, n. \end{aligned} \quad (2.5)$$

한편 Mercer (1909)의 조건에 의하면 $\boldsymbol{\phi}(\mathbf{x}_{1s})' \boldsymbol{\phi}(\mathbf{x}_{1t}) = K(\mathbf{x}_{1s}, \mathbf{x}_{1t})$ 가 성립한다. 여기서 $K(\cdot, \cdot)$ 는 커널함수이다. 따라서 커널함수 K 만을 사용하고 특징함수 $\boldsymbol{\phi}$ 가 구체적으로 무엇인지를 알 필요가 없다. 일반적인 커널함수는 다음과 같은 다항커널함수와 RBF (radial basis function) 커널함수이다.

$$\text{다항커널함수} : K(\mathbf{x}_s, \mathbf{x}_t) = (\mathbf{x}'_s \mathbf{x}_t + 1)^p, \quad \text{RBF} : K(\mathbf{x}_s, \mathbf{x}_t) = e^{-\frac{\|\mathbf{x}_s - \mathbf{x}_t\|^2}{s^2}}$$

여기서 p 와 s^2 은 커널모수이다.

커널함수를 사용하면, 식 (2.3)를 최소화하여 $\boldsymbol{\omega}_1$ 와 b_1 의 추정값들을 구하는 문제는 식 (2.5)에 의해 식 (2.6)을 만족하는 $\boldsymbol{\alpha}_1$ 과 b_1 을 구하는 문제로 바뀌게 된다.

$$\mathbf{y} - K_1 \boldsymbol{\alpha}_1 - b_1 \mathbf{1} - \frac{1}{C_1} W \boldsymbol{\alpha}_1 = \mathbf{0}, \boldsymbol{\alpha}'_1 \mathbf{1} = 0, \quad (2.6)$$

여기서 $\mathbf{y} = (y_1, y_2, \dots, y_n)'$, 커널행렬 K_1 은 원소가 $K(\mathbf{x}_{1s}, \mathbf{x}_{1t}), s, t = 1, \dots, n$ 인 $n \times n$ 행렬이며 W 는 σ_t^2 로 이루어진 대각행렬이다. 따라서 라그랑제 배수 $\boldsymbol{\alpha}_1$ 과 b_1 의 추정값은 다음 식 (2.7)과 같은 선형방정식의 해가 된다.

$$\begin{pmatrix} K_1 + W/C_1 & \mathbf{1} \\ \mathbf{1}' & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha}_1 \\ b_1 \end{pmatrix} = \begin{pmatrix} \mathbf{y} \\ 0 \end{pmatrix}, \quad (2.7)$$

한편, μ_t 의 추정값은 $\hat{\mu}_t = \mathbf{k}_{1t} \boldsymbol{\alpha}_1 + b_1$ 로 구해진다. 여기서 \mathbf{k}_{1t} 는 커널행렬 K_1 의 t 번째 행벡터를 나타낸다.

2.2. 변동성의 추정

한편 분산은 음수가 될 수 없기 때문에 σ_t^2 대신 로그분산 $g_t = \log \sigma_t^2$ 를 사용하여 식 (2.2)의 분산방정식 대신 다음과 같은 분산방정식을 고려한다.

$$\log \sigma_t^2 = f_3(\varepsilon_{t-1}^2, \sigma_{t-1}^2) \quad (2.8)$$

여기서 f_3 은 $\mathbf{x}_{2t} = (\varepsilon_{t-1}^2, \sigma_{t-1}^2)$ 의 비선형 함수이다. 우리는 비가중 LS-SVM (Suykens과 Vanderwalle, 1999)을 이용하여 $g_t = g(\mathbf{x}_{2t})$ 의 추정값을 구한다.

커널행렬 K_2 는 원소가 $K(\mathbf{x}_{2s}, \mathbf{x}_{2t}), s, t = 1, \dots, n$ 인 $n \times n$ 행렬을 나타내고 \mathbf{k}_{2t} 는 K_2 의 t 번째 행 벡터를 나타낼 때 $g_t = g(\mathbf{x}_{2t})$ 의 추정값 $\hat{g}_t = \mathbf{k}_{2t}\boldsymbol{\alpha}_2 + b_2$ 의 $\boldsymbol{\alpha}_2$ 와 b_2 는 KKT 조건 (Kuhn과 Tucker, 1951)에 의하여 다음 조건을 만족한다.

$$\mathbf{g} - K_2\boldsymbol{\alpha}_2 - \mathbf{1}b_2 - \frac{1}{C_2}\boldsymbol{\alpha}_2 = \mathbf{0}, \quad \boldsymbol{\alpha}'_2\mathbf{1} = 0, \quad (2.9)$$

여기서 $C_2 > 0$ 는 벌칙상수이다. 라그랑제 배수 $\boldsymbol{\alpha}_2$ 와 b_2 의 추정값은 다음의 선형방정식의 해로서 구해진다.

$$\begin{pmatrix} K_2 + I/C_2 & \mathbf{1} \\ \mathbf{1}' & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha}_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} \mathbf{g} \\ 0 \end{pmatrix}. \quad (2.10)$$

따라서 변동성 (σ_t)의 추정값은 $\hat{\sigma}_t = \exp(\hat{g}_t/2) = \exp((K(\mathbf{x}_{2t}, \mathbf{x}_2)\boldsymbol{\alpha}_2 + b_2)/2)$ 로 구해지고, 이 경우 $\mathbf{x}_2 = (\mathbf{x}'_{21}, \mathbf{x}'_{22}, \dots, \mathbf{x}'_{2n})'$ 이다.

2.3. 평균과 변동성의 추정 알고리즘

일반적으로 변동성 σ_t 의 입력자료는 주어지지 않으므로 다음과 같은 반복 알고리즘을 이용하여 평균 μ_t 와 변동성 σ_t 를 추정한다.

[단계1] 입력변수를 $\mathbf{x}_{1t} = y_{t-1}, W = I$ 로 두고 $\hat{\mu}_t = K(\mathbf{x}_{1t}, \mathbf{x}_1)\boldsymbol{\alpha}_1 + b_1$ 와 $\hat{\varepsilon}_t = y_t - \hat{\mu}_t$ 를 식 (2.7)을 이용하여 구한다. 여기서, $\mathbf{x}_1 = (x_{11}, x_{12}, \dots, x_{1n})'$ 이다.

[단계2] 출력변수를 $\log \hat{\sigma}_t^2$, 입력변수를 $\mathbf{x}_{2t} = (\hat{\varepsilon}_{t-1}^2, \hat{\sigma}_{t-1}^2)'$ 로 놓고, 식 (2.10)을 이용하여 $\hat{g}_t = K(\mathbf{x}_{2t}, \mathbf{x}_2) + b_2$ 를 구한다. 여기서, $\hat{\sigma}_t^2 = \{\sum_{k=0}^{m-1} (y_{t-k} - \hat{\mu}_{t-k})^2\}/m$ 이다.

[단계3] 입력변수를 $\mathbf{x}_{1t} = (y_{t-1}, \hat{\sigma}_t)$ 로 놓고 $\hat{\mu}_t = K(\mathbf{x}_{1t}, \mathbf{x}_1)\boldsymbol{\alpha}_1 + b_1$ 와 $\hat{\varepsilon}_t = y_t - \hat{\mu}_t$ 를 식 (2.7)을 이용하여 구한다. 여기서, $\hat{\sigma}_t = \exp(\hat{g}_t/2)$ 이다.

[단계4] 출력변수를 \hat{g}_t , 입력변수를 $\mathbf{x}_{2t} = (\hat{\varepsilon}_{t-1}^2, \hat{\sigma}_{t-1}^2)'$ 로 놓고, 식 (2.10)을 이용하여 $\hat{g}_t = K(\mathbf{x}_{2t}, \mathbf{x}_2) + b_2$ 를 구한다. 여기서, $\hat{\sigma}_t^2 = \exp(\hat{g}_t)$ 이다.

[단계5] $\{\sum_{t=1}^n (\hat{g}_t^{(l)} - \hat{g}_t^{(l-1)})^2\}/n$ 이 수렴할 때까지 [단계3]과 [단계4]를 반복한다.

3. 모형선택

비선형 GARCH-M모형의 추정 알고리즘의 성능은 벌칙상수와 커널모수에 의해 결정된다. 벌칙상수 (C_1, C_2)와 커널행렬 K_1 과 K_2 에 포함된 커널모수 (s_1^2, s_2^2)의 최적값을 구하기 위해서 우리는 다음과 같

은 교차타당성 (cross validation, CV) 함수를 고려한다.

$$CV(\boldsymbol{\theta}) = \frac{1}{n} \sum_{t=1}^n \frac{1}{\hat{\sigma}_t^2} (y_t - \hat{\mu}_t^{(-t)})^2 \tag{3.1}$$

여기서 $\boldsymbol{\theta}$ 는 벌칙상수와 커널모수의 집합 (C_1, C_2, s_1^2, s_2^2) 이고, $\hat{\sigma}_t^2$ 는 전체자료를 이용하여 구한 σ_t^2 의 추정값이며, $\hat{\mu}_t^{(-t)}$ 는 t 번째 자료를 뺀 나머지 자료를 이용하여 구한 μ_t 의 추정값이다. 벌칙상수와 커널모수의 값들의 각 조합에 대하여 n 번의 추정값을 구하여야 하므로 실제로 교차타당성함수를 이용하여 벌칙상수와 커널모수의 최적값을 구하는 것은 계산적으로 부하가 크다. 따라서 leaving-out-one lemma (Kimeldorf 와 Wahba, 1971)와 1차 테일러 급수 (first-order Taylor expansion)를 이용하면

$$(y_t - \hat{\mu}_t^{(-t)}) - (y_t - \hat{\mu}_t) = \hat{\mu}_t - \hat{\mu}_t^{(-t)} \approx \frac{\partial \hat{\mu}_t}{\partial y_t} (y_t - \hat{\mu}_t^{(-t)})$$

이고, $(y_t - \hat{\mu}_t^{(-t)}) \approx (y_t - \hat{\mu}_t)/(1 - \partial \hat{\mu}_t/\partial y_t) = (y_t - \hat{\mu}_t)/(1 - s_{tt})$ 가 된다.

여기서 s_{tt} 는 식 (2.7)에서 유도되는 $\hat{\boldsymbol{\mu}} = S\mathbf{y}$ 를 만족하는 모자행렬 (hat matrix) S 의 t 번째 대각원소이고 S 는 다음과 같이 구해진다.

$$S = (K_1, \mathbf{1})H_{.1}, \begin{pmatrix} H_{.1}, H_{.2} \\ \mathbf{1}' & 0 \end{pmatrix}^{-1}$$

여기서 $H_{.1}$ 은 $(n+1) \times n$ 행렬이다. 따라서 식 (3.1)의 교차타당성 (CV) 함수는 다음과 같이 근사된다.

$$CV(\boldsymbol{\theta}) \approx \frac{1}{n} \sum_{t=1}^n \frac{1}{\hat{\sigma}_t^2} \left(\frac{y_t - \hat{\mu}_t}{1 - s_{tt}} \right)^2 \tag{3.2}$$

식 (3.2)에서 s_{tt} 대신 $trace(S)/n$ 을 사용하면 일반화 교차타당성 (generalized cross validation, GCV) 함수는 다음과 같이 구해진다.

$$GCV(\boldsymbol{\theta}) = \frac{n}{(n - trace(S))^2} \sum_{t=1}^n \frac{1}{\hat{\sigma}_t^2} (y_t - \hat{\mu}_t)^2 \tag{3.3}$$

4. 실제자료 분석

이절에서는 1990년 1월부터 1998년 12월까지 조사된 S&P500지수 (US Standard & Poor's 500 index)와 30년 만기 미국재무성채권 (30-years US Treasury bonds) 자료에 대하여 제안된 비선형 GARCH(1,1)-M모형을 적용하여 선형 GARCH(1,1)모형 및 선형 GARCH(1,1)-M모형과 성능을 비교한다. 주어진 자료는 각각 2212개의 일일 로그수익률 (log-return)과 실현변동성 (realized volatility)으로 구성되어 있다. 그림 4.1은 S&P500지수 자료와 미국재무성채권 자료의 로그수익률 (y_t)과 실현변동성 (σ_t)을 보여준다. 그림 4.1의 위는 S&P500지수 자료의 로그수익률과 실현변동성, 아래는 미국재무성채권 자료의 로그수익률과 실현변동성을 보여준다. 추정의 성능을 계량화하기 위해서 우리는 다음과 같이 로그수익률과 실현변동성을 이용한 MAE과 MSE를 사용한다.

$$MAE_\mu = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{\mu}_t|, \quad MSE_\mu = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{\mu}_t)^2,$$

$$MAE_\sigma = \frac{1}{n} \sum_{t=1}^n |\sigma_t - \hat{\sigma}_t|, \quad MSE_\sigma = \frac{1}{n} \sum_{t=1}^n (\sigma_t - \hat{\sigma}_t)^2$$

선형 GARCH(1,1)모형으로, S&P500지수 자료의 변동성은 $\hat{\sigma}_t^2 = 2.3839 + 0.0730\hat{\varepsilon}_{t-1}^2 + 0.9163\hat{\sigma}_{t-1}^2$ 로 추정되고, 미국재무성채권 자료의 변동성은 $\hat{\sigma}_t^2 = 0.7005 + 0.0160\hat{\varepsilon}_{t-1}^2 + 0.9735\hat{\sigma}_{t-1}^2$ 로 추정된다. 선형 GARCH(1,1)-M모형으로, S&P500지수 자료의 평균은 $\hat{\mu}_t = -1.9457 - 0.0648y_{t-1} + 0.2081\hat{\sigma}_t$ 로 추정되고, 미국재무성채권 자료의 평균은 $\hat{\mu}_t = 1.6926 - 0.0248y_{t-1} - 0.1762\hat{\sigma}_t$ 로 추정된다. 또한 S&P500지수 자료의 변동성은 $\hat{\sigma}_t^2 = 2.2420 + 0.0682\hat{\varepsilon}_{t-1}^2 + 0.9215\hat{\sigma}_{t-1}^2$ 로 추정되고, 미국재무성채권 자료의 변동성은 $\hat{\sigma}_t^2 = 0.6996 + 0.0160\hat{\varepsilon}_{t-1}^2 + 0.9735\hat{\sigma}_{t-1}^2$ 로 추정된다.

추정 알고리즘 [단계2]의 입력변수 \mathbf{x}_{2t} 에 사용된 $\hat{\sigma}_t^2$ 는 $(y_t - \hat{\mu}_t)^2$ 의 20일 사전이동평균을 사용하였으며 μ_t 와 σ_t^2 추정을 위하여 RBF 커널함수를 이용한다. μ_t 와 σ_t^2 추정을 위한 벌칙상수 (C_1, C_2)와 커널모수 (s_1^2, s_2^2)는 3절의 일반화 교차타당성 (GCV) 함수 (3.3)을 이용해서 구해진다. S&P500지수 자료에서 (μ_t, σ_t^2) 추정의 벌칙상수와 커널모수는 $(C_1, s_1^2) = (100, 12)$, $(C_2, s_2^2) = (100, 1)$ 로 구해진다. 또한 미국재무성채권 자료에서 (μ_t, σ_t^2) 추정의 벌칙상수와 커널모수는 $(C_1, s_1^2) = (100, 8)$, $(C_2, s_2^2) = (100, 1)$ 로 구해진다.

선형 GARCH(1,1)모형, 선형 GARCH(1,1)-M모형 그리고 제안된 비선형 GARCH(1,1)-M모형의 모형식은 표 4.1에 주어지고, 각 모형에 의한 평균 및 변동성 추정의 MAE와 MSE값은 표 4.2와 표 4.3에 각각 주어지고 있다. 우리는 두 자료에 대해서 MAE와 MSE값 둘 다 제안된 비선형 GARCH(1,1)-M모형이 선형 GARCH(1,1)모형과 선형 GARCH(1,1)-M모형보다 대체적으로 작은 값을 나타냄을 알 수 있다. 즉 주어진 두 자료에 대하여 제안된 비선형 GARCH(1,1)-M모형이 선형 GARCH(1,1) 및 선형 GARCH(1,1)-M모형보다 우수한 성능을 보임을 알 수 있다. 따라서 S&P500지수와 30년 만기 미국재무성채권 자료의 수익률의 평균과 변동성예측을 통하여 금융시계열자료의 평균과 변동성 예측을 위해 주로 사용되는 GARCH-M모형의 추정을 위하여 본 논문에서 제안한 최소제곱 서포트벡터기계 기반방법이 기존의 방법들 보다 더 나은 추정능력을 가진다는 것을 확인할 수 있었다.

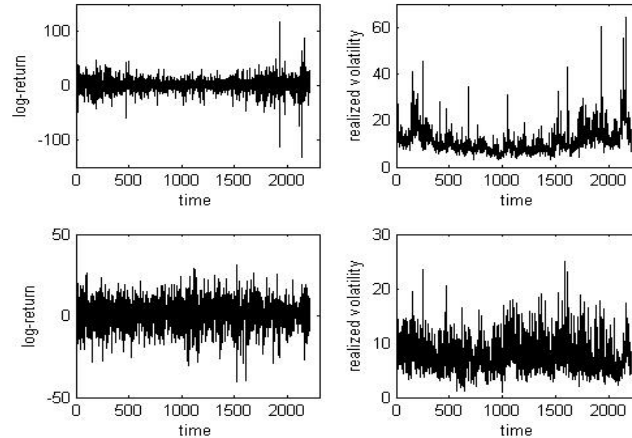


그림 4.1 실제자료의 로그수익률 및 실현변동성 (위: S&P500지수자료, 아래: 30년 만기 미국재무성채권 자료)

표 4.1 적용 모형식

모형	모형식
선형 GARCH(1,1)	$y_t = \mu + \varepsilon_t, \sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$
선형 GARCH(1,1)-M	$y_t = \gamma_0 + \gamma_1 y_{t-1} + \gamma_2 \sigma_t + \varepsilon_t$ $\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2$
비선형 GARCH(1,1)-M	$y_t = f_1(y_{t-1}, \sigma_t) + \varepsilon_t$ $\sigma_t^2 = f_2(\varepsilon_{t-1}^2, \sigma_{t-1}^2)$

표 4.2 실제자료의 평균 추정의 성능비교

	S&P500지수 자료		미국재무성채권 자료	
	MAE _μ	MSE _μ	MAE _μ	MSE _μ
선형 GARCH(1,1)-M	9.4095	181.8694	6.0778	66.4699
비선형 GARCH(1,1)-M	9.0948	167.8461	5.9341	63.2494

표 4.3 실제자료의 변동성 추정의 성능비교

	S&P500지수 자료		미국재무성채권 자료	
	MAE _σ	MSE _σ	MAE _σ	MSE _σ
선형 GARCH(1,1)	3.0589	19.6339	2.2656	8.1864
선형 GARCH(1,1)-M	3.0367	18.9540	2.2641	8.1791
비선형 GARCH(1,1)-M	2.7846	17.4165	2.1508	7.7286

참고문헌

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

Domowitz, I. and Hakkio, C. S. (1985). Conditional variance and the risk premium in the foreign exchange market. *Journal of International Economics*, **18**, 47-66.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**, 987-1007.

Hwang, C. (2007). Kernel machine for Poisson regression. *Journal of Korean Data & Information Science Society*, **18**, 767-772.

Hwang, C. and Shin, S. (2010). Estimating GARCH models using kernel machine learning. *Journal of Korean Data & Information Science Society*, **21**, 419-425.

Kimeldorf, G. S. and Wahba, G. (1971). Some results on Tchebycheffian spline functions. *Journal of Mathematical Analysis and its Applications*, **33**, 82-95.

Kuhn, H. W. and Tucker, A. W. (1951). Nonlinear programming. *Proceedings of 2nd Berkeley Symposium*, University of California Press, Berkeley, 481-492.

Mercer, J. (1909). Function of positive and negative type and their connection with theory of integral equations. *Philosophical Transactions of Royal Society*, **A**, 415-446.

Najand, M. and Yung, K. (1991). A GARCH examination of the relationship between volume and price variability in futures markets. *The Journal of Futures Markets*, **11**, 613-621.

Perez Cruz, F., Afonso Rodriguez, J. A. and Giner, J. (2003). Estimating GARCH models using support vector machines. *Quantitative Finance*, **3**, 163-172.

Shim, J. and Hwang, C. (2003). Prediction intervals for LS-SVM regression using the bootstrap. *Journal of Korean Data & Information Science Society*, **14**, 337-343.

Shim, J. and Lee, J. T. (2009). Kernel method for autoregressive data. *Journal of Korean Data & Information Science Society*, **20**, 467-472.

Suykens, J. A. K. and Vanderwalle, J. (1999). Least square support vector machine classifier. *Neural Processing Letters*, **9**, 293-300.

- Suykens, J. A. K., De Brabanter, J., Lukas, L. and Vandewalle, J. (2002). Weighted least squares support vector machines: Robustness and sparse approximation. *Neurocomputing*, **66**, 85-105
- Vapnik, V. N. (1998). *Statistical Learning Theory*, John Wiley, New York.

Estimation of nonlinear GARCH-M model[†]

Jooyong Shim¹ · Jangtaek Lee²

¹Department of Applied Statistics, Catholic University of Daegu

²Department of Statistics, Dankook University

Received 13 May 2010, revised 12 August 2010, accepted 28 August 2010

Abstract

Least squares support vector machine (LS-SVM) is a kernel trick gaining a lot of popularities in the regression and classification problems. We use LS-SVM to propose a iterative algorithm for a nonlinear generalized autoregressive conditional heteroscedasticity model in the mean (GARCH-M) model to estimate the mean and the conditional volatility of stock market returns. The proposed method combines a weighted LS-SVM for the mean and unweighted LS-SVM for the conditional volatility. In this paper, we show that nonlinear GARCH-M models have a higher performance than the linear GARCH model and the linear GARCH-M model via real data estimations.

Keywords: Generalized autoregressive conditional heteroscedasticity model, generalized autoregressive conditional heteroscedasticity model in the mean model, generalized cross validation, least squares support vector machine, weighted least squares support vector machine.

[†] This research was supported by the National IT Industry Promotion Agency (NIPA) under the program of Software Engineering Technologies Development.

¹ Adjunct professor, Department of Applied Statistics, Catholic University of Daegu, Gyungbuk 702-701, Korea.

² Corresponding author: Professor, Department of Statistics, Dankook University, Yongin 448-701, Korea. E-mail: jtlee@dankook.ac.kr