Journal of the Korean Data & Information Science Society
2010, **21**(3), 419–425

한국데이터정보과학회지

커널기계 기법을 이용한 일반화 이분산자기회귀모형 추정†

황창하[1] · 신사임[2]

[12]단국대학교 정보통계학과

접수 2010년 3월 22일, 수정 2010년 5월 8일, 게재확정 2010년 5월 17일

요 약

커널기계 기법은 최근 대용량 또는 고차원 비선형 자료를 분석하는 방법으로 인기를 많이 얻고 있다. 본 논문에서는 주식시장 수익률의 조건부 변동성을 예측하기 위한 일반화 이분산자기회귀모형을 추정하기 위해 커널기계 기법을 사용한다. 일반화 이분산자기회귀모형은 자료가 정규분포를 따른다고 가정한 후 주로 최대우도법을 사용하여 추정된다. 본 논문에서는 꼬리가 두꺼운 분포를 갖는 금융시계열자료의 변동성을 추정할 때 커널기계 기법이 최대우도법과 서포트벡터기계 보다 더 정확한 예측능력을 가진다는 것을 보이고자 한다.

주요용어: 서포트벡터기계, 일반화근사교차타당성, 일반화 이분산자기회귀모형, 커널기계, 코스피.

1. 서론

주식시장에서의 주가가 급락과 등락을 반복하는 현상이 관측될 때 많은 사람들이 변동성에 대해 관심을 가지고 분석을 하게 된다. 금융시계열의 변동성에 대한 정교한 추정과 예측은 금융모형분석 또는 이론검정에 있어서 중요할 뿐만 아니라 실무에서 선물, 옵션 등의 파생상품을 거래하는 투자자에게도 매우 중요하다. 변동성군집 또는 두꺼운 꼬리의 분포를 갖는 금융시계열을 조건부 분산의 관점에서 모형화하기 위하여 Engle (1982)은 ARCH (autoregressive conditional heteroscedasticity) 모형을 이용하여 시간가변적 변동성을 포착하는데 처음으로 성공하였다. ARCH모형은 자산수익률의 변동성의 특징을 표현하기 위해서 GARCH, EGARCH, IGARCH, TGARCH 등과 같은 조건부 분산모형으로 확대되었다. 이런 GARCH 모형들은 실제 금융시계열자료를 분석하기위해 널리 사용되고 있다. GARCH 모형 중 가장 대표적인 모형은 GARCH(1,1)이다. 자세한 내용은 김명직과 장국현 (2002) 또는 Perez-Cruz 등 (2003)에 설명되어 있다. GARCH 모형들의 추정을 위해 기본적으로 최대우도법을 사용한다. 그러나 더 정확한 추정결과를 얻기 위해 추정방법에 대한 연구의 필요성이 대두되고 있다. 최근 Perez-Cruz 등 (2003)은 정규분포를 따르지 않는 자료에 적용할 수 있는 GARCH 모형을 추정하기위해 SVM (support vector machine)을 제안하였다. 본 논문에서는 GARCH 모형을 추정하기 위해 커널기계 기법을 제안하고 기존의 최대우도법, SVM과 비교하고자한다. 통계학에서 커널기계 기법을 사용한 최근 논문들은 Hwang (2007, 2008), Shim과 Seok (2008), Shim 등 (2009) 등이 있다.

---

† 이 연구는 2009년도 단국대학교 대학원 연구보조장학금의 지원으로 이루어진 것임.
[1] 교신저자: (448-701) 경기도 용인시 수지구 죽전동 126번지, 단국대학교 정보통계학과, 교수.
E-mail: chwang@dankook.ac.kr
[2] (448-701) 경기도 용인시 수지구 죽전동 126번지, 단국대학교 정보통계학과, 석사과정.

만약 y_t가 GARCH (1,1) 모형을 따른다면 평균방정식과 분산방정식은 다음과 같이 정의된다.

$$\text{평균방정식:}\ y_t = \mu + \sigma_t \varepsilon_t, t = 1, \cdots, n. \tag{1.1}$$
$$\text{분산방정식:}\ \sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2.$$

여기서 ε_t는 일반적으로 $N(0,1)$을 따른다고 가정한다. 따라서 오차항 $\sigma_t\varepsilon_t$은 $N(0,\sigma_t^2)$을 따른다. 지금부터 일반성을 잃지 않고 $\mu = 0$을 가정한다. 분산방정식의 모수 ω, α, β는 조건부 분산이 양수이어야 하므로 $\omega > 0,\ \alpha, \beta \geq 0$을 만족해야 한다. 또한 GARCH (1,1) 모형에서 분산이 양이고 정상성 조건을 만족하기 위한 조건은 $\alpha + \beta < 1$이어야 한다. 오차항을 정규분포로 가정하는 GARCH (1,1) 모형은 꼬리가 두꺼운 분포를 따르는 수익률 시계열자료를 충분히 다 설명할 수 없는 경우가 많다. 따라서 Perez-Cruz 등 (2003)은 오차항이 정규분포를 따르지 않는 GARCH (1,1) 모형의 추정법으로 SVM 기법을 제안하였다. SVM은 관측가능한 종속변수와 독립변수를 필요로 하는데, 종속변수는 σ_t^2 이고 독립변수들은 $y_{t-1}^2, \sigma_{t-1}^2$이다. 그리고 σ_t^2을 y_t^2의 5일 이동평균인 $\hat{\sigma}_t^2 = \frac{1}{5}\sum_{k=0}^4 y_{t-k}^2$으로 측정한다.

본 논문은 다음과 같이 구성되어 있다. 2절에서는 GARCH 모형의 추정을 위한 커널기계에 대해서 3절에서는 커널기계의 구조를 결정하는 벌칙상수와 커널모수를 결정하는 방법에 대해서 기술하였다. 4절과 5절에서는 각각 모의실험과 KOSPI 자료분석 에 대한 결과를 설명한다. 마지막 6절은 결론을 기술하고 있다.

2. GARCH (1,1) 모형추정을 위한 커널기계

커널기계 기법은 비선형 문제를 풀기위해 입력공간의 입력벡터를 더 높은 차원의 특징공간 (feature space)으로 사상 (mapping)하여 선형모형을 이용하는 방법이다. 분산함수의 벌칙로그우도 (penalized log-likelihood)함수를 최소화하여 분산함수를 추정하는 방법을 설명하면 다음과 같다.

주어진 시계열자료 $\{y_t\}_{t=1}^n$에 대해 다음과 같은 GARCH (1,1) 모형을 고려한다.

$$\text{평균방정식:}\ y_t = \sigma_t \varepsilon_t, t = 1, 2, \cdots, n. \tag{2.1}$$
$$\text{분산방정식:}\ \sigma_t^2 = f(y_{t-1}^2, \sigma_{t-1}^2).$$

여기서 $f(\cdot,\cdot)$는 미지의 비선형 함수이다. 만약 오차항 ε_t가 표준정규분포 $N(0,1)$를 따른다고 가정하면 음로그우도 (negative log-likelihood)함수는 상수항을 생략한 후 다음과 같이 표현된다.

$$L(\sigma^2) = \frac{1}{n}\sum_{t=1}^n \left[\frac{y_t^2}{\sigma_t^2} + \log \sigma_t^2\right]. \tag{2.2}$$

일반적으로 오차항이 정규분포를 따르는 것과 상관없이 식 (2.2)는 분산함수를 추정할 때 목적함수로 자주 사용된다. 자세한 내용은 Audrino와 Bühlmann (2009)에 설명되어 있다. 식 (2.2)는 감마분포로부터 독립적으로 추출된 y_t^2들의 음우도함수를 나타내는 것으로 볼 수 있다. 그리고 식 (2.2)를 최소화하여 구한 추정량은 상당히 로버스트 (robust)한 것으로 알려져 있다. 자세한 내용은 Juutilainen과 Rüning (2006)에 설명되어 있다. 이런 이유로 오차항이 정규분포를 따른 것과 상관없이 본 논문에서도 식 (2.2)를 목적함수로 사용하고 커널기계 기법을 적용하여 이 함수를 최소화하는 분산함수 σ_t^2를 구하고자한다. 한편 분산은 음수가 될 수 없기 때문에 σ_t^2 대신에 로그분산 $g(\boldsymbol{x}_t) = \log \sigma_t^2$를 사용하면 식 (2.2)는 다음과 같이 표현된다. 이때 σ_t^2은 y_{t-1}^2과 σ_{t-1}^2의 함수이기 때문에 $\boldsymbol{x}_t = (y_{t-1}^2, \sigma_{t-1}^2)$를 사용하

여 $g(\boldsymbol{x}_t) = \log \sigma_t^2$로 나타낸다.

$$L(\boldsymbol{g}) = \frac{1}{n}\sum_{t=1}^{n}\left[y_t^2 e^{-g(\boldsymbol{x}_t)} + g(\boldsymbol{x}_t)\right]. \tag{2.3}$$

커널기계 기법은 로그분산 $g(\boldsymbol{x}_t) = \log \sigma_t^2$을 고차원 특징공간에서의 선형모형 $g(\boldsymbol{x}_t) = \boldsymbol{\omega}^T \boldsymbol{\phi}(\boldsymbol{x}_t) + b$를 사용하여 추정한다. 이때 특징함수 $\boldsymbol{\phi}(\cdot) : R^d \to R^{d_f}$는 입력공간을 더 높은 차원의 특징공간으로 사상시킨다. 이제 $\boldsymbol{\omega}$ 와 b의 추정값들을 구하기 위하여 다음과 같이 수정된 벌칙음로그우도함수

$$L(\boldsymbol{\omega}, b) = \sum_{t=1}^{n}\left[y_t^2 e^{-\boldsymbol{\omega}^t \boldsymbol{\phi}(\boldsymbol{x}_t) - b} + \boldsymbol{\omega}^t \boldsymbol{\phi}(\boldsymbol{x}_t) + b\right] + \frac{\lambda}{2}\|\boldsymbol{\omega}\|^2 \tag{2.4}$$

를 최소화한다. 여기서 λ는 자료의 적합 (fitting)과 함수의 평탄성 (flatness)을 조절하는 벌칙상수이다. 한편 Mercer (1909)의 조건에 의하면 $\boldsymbol{\phi}(\boldsymbol{x}_s)^T \boldsymbol{\phi}(\boldsymbol{x}_t) = K(\boldsymbol{x}_s, \boldsymbol{x}_t)$가 성립한다. 여기서 $K(\cdot,\cdot)$는 커널함수이다. 따라서 커널함수 K만을 사용하고 특징함수 $\boldsymbol{\phi}$가 구체적으로 무엇인지를 알 필요가 없다. 일반적인 커널함수는 다음과 같은 다항커널함수와 RBF커널함수이다. 본 논문에서는 RBF커널함수를 사용한다.

다항커널함수: $K(\boldsymbol{x}_s, \boldsymbol{x}_t) = (\boldsymbol{x}_s^T \boldsymbol{x}_t + 1)^p$, RBF커널함수: $K(\boldsymbol{x}_s, \boldsymbol{x}_t) = e^{-\|\boldsymbol{x}_s - \boldsymbol{x}_t\|^2/2\sigma^2}$

여기서 p와 σ^2은 커널모수이다. 커널함수를 사용하면 식 (2.4)를 최소화하여 $\boldsymbol{\omega}$와 b의 추정값들을 구하는 문제는 다음의 식 (2.5)를 최소화하여 $\boldsymbol{\alpha}$와 b를 구하는 문제로 바뀌게 된다.

$$L(\boldsymbol{\alpha}, b) = \boldsymbol{z}^T \boldsymbol{e}^{-\boldsymbol{K}\boldsymbol{\alpha} - b\mathbf{1}_n} + \mathbf{1}_n^T(\boldsymbol{K}\boldsymbol{\alpha} + b\mathbf{1}_n) + \frac{1}{2}\boldsymbol{\alpha}^T \boldsymbol{K}\boldsymbol{\alpha}. \tag{2.5}$$

여기서 $\boldsymbol{z} = (y_1^2, \cdots, y_n^2)^T$, $\mathbf{1}_n = (1, \cdots, 1)^T$, 행렬 $\boldsymbol{K}$는 원소가 $K(\boldsymbol{x}_s, \boldsymbol{x}_t), s, t = 1, \cdots, n$인 $n \times n$ 행렬이다. 그리고 지수함수 $\boldsymbol{e}$는 대응원소별 (componentwise) 지수함수이다. 식 (2.5)는 마치 형상모수 (shape parameter)가 1이고 척도모수 (scale parameter)가 $\exp(g_t) = \exp(\boldsymbol{k}_t\boldsymbol{\alpha} + b)$인 감마분포로부터 독립적으로 추출된 z_t들의 벌칙음우도함수를 나타내는 것으로 볼 수 있다. 여기서 $\boldsymbol{k}_t$는 커널행렬 $\boldsymbol{K}$ 의 t번째 행벡터를 나타낸다.

Newton-Raphson 방법을 사용하여 식 (2.5)를 최소화하는 $\boldsymbol{\alpha}$와 b의 추정량을 반복적으로 구할 수 있는데 $t+1$번째 반복단계에서의 $\boldsymbol{\alpha}$와 b의 추정량은 다음과 같이 최신화 (update)된다.

$$\begin{pmatrix} \boldsymbol{\alpha}^{(t+1)} \\ b^{(t+1)} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}^{(t)} \\ b^{(t)} \end{pmatrix} - \boldsymbol{H}(\boldsymbol{\alpha}^{(t)}, b^{(t)})^{-1} \boldsymbol{G}(alpha^{(t)}, b^{(t)}). \tag{2.6}$$

여기서 식 (2.5)의 일차미분인 $(n+1) \times 1$ 벡터 $\boldsymbol{G}(\boldsymbol{\alpha}^{(t)}, b^{(t)})$와 이차미분인 $(n+1) \times (n+1)$ 행렬 $\boldsymbol{H}(\boldsymbol{\alpha}^{(t)}, b^{(t)})$는 다음과 같이 주어진다.

$$\boldsymbol{G}(\boldsymbol{\alpha}^{(t)}, b^{(t)}) = \begin{pmatrix} -\boldsymbol{K}\boldsymbol{Z}\boldsymbol{e}^{-\boldsymbol{K}\boldsymbol{\alpha}^{(t)} - b^{(t)}\mathbf{1}_n} + \boldsymbol{K}\mathbf{1}_n + \lambda \boldsymbol{K}\boldsymbol{\alpha}^{(t)} \\ -\mathbf{1}_n^T \boldsymbol{Z}\boldsymbol{e}^{-\boldsymbol{K}\boldsymbol{\alpha}^{(t)} - b^{(t)}\mathbf{1}_n} + n \end{pmatrix}. \tag{2.7}$$

$$\boldsymbol{H}(\boldsymbol{\alpha}^{(t)}, b^{(t)}) = \begin{pmatrix} \boldsymbol{K}\boldsymbol{Z}\boldsymbol{E}\boldsymbol{K} + \lambda \boldsymbol{K} & \boldsymbol{K}\boldsymbol{Z}\boldsymbol{E}\mathbf{1}_n \\ \mathbf{1}_n^T \boldsymbol{Z}\boldsymbol{E}\boldsymbol{K} & \mathbf{1}_n^T \boldsymbol{Z}\boldsymbol{E}\mathbf{1}_n \end{pmatrix}. \tag{2.8}$$

이때 행렬 $\boldsymbol{Z}$는 대각원소들이 $y_1^2, \cdots, y_n^2$인 대각행렬이고, 즉 $\boldsymbol{Z} = \mathrm{diag}(z_1, \cdots, z_n)$이고, 행렬 $\boldsymbol{E}$는 $n \times 1$ 벡터 $\boldsymbol{e}^{-\boldsymbol{K}\boldsymbol{\alpha}^{(t)} - b^{(t)}\mathbf{1}_n}$의 원소들이 대각원소를 구성하는 대각행렬이다.

3. 모형선택

GARCH 모형의 추정을 위한 커널기계의 구조는 벌칙상수 λ와 커널모수 σ^2에 의해 결정된다. 이 모수들은 Xiang과 Wahba (1996)에 의해 제안된 일반화근사교차타당성 (generalized apprioximate cross validation, GACV) 기법을 시용하여 결정할 수 있는데, GACV 함수를 유도하는 과정을 간단하게 설명하면 다음과 같다. 먼저 Leave-One-Out 교차타당성 (cross validation, CV) 함수는 다음과 같이 정의된다.

$$CV(\boldsymbol{\theta}) = \frac{1}{n}\sum_{t=1}^{n}(z_t e^{-\hat{g}_t^{(-t)}} + \hat{g}_t). \tag{3.1}$$

여기서 $\boldsymbol{\theta} = (\lambda, \sigma^2)$, $\hat{g}_t$는 모든 자료를 사용하여 구한 g_t의 추정값이고 $\hat{g}_t^{(-t)}$는 t번째 관측치를 제거한 자료를 사용하여 구한 g_t의 추정값이다. 교차타당성 함수를 사용해서 모수를 선택하는 것은 계산량이 많다. 따라서 다음과 같은 근사교차타당성 (approximate cross validation, ACV)을 사용한다.

$$ACV(\boldsymbol{\theta}) = \frac{1}{n}\sum_{t=1}^{n}(z_t e^{-\hat{g}_t} + \hat{g}_t) + \frac{1}{n}\sum_{t=1}^{n}\frac{h_{tt}e^{-\hat{g}_t}z_t(z_t - e^{\hat{g}_t})}{1 - h_{tt}e^{\hat{g}_t}}. \tag{3.2}$$

여기서 h_{tt}는 행렬 $\boldsymbol{H} = (\boldsymbol{W} + \lambda\boldsymbol{I})^{-1}\boldsymbol{V}$의 i번째 대각원소이다. 이때 행렬 $\boldsymbol{W}$와 $\boldsymbol{V}$는 각각 $\boldsymbol{W} = \mathrm{diag}(z_i e^{-\hat{g}_1}, \cdots, z_n e^{-\hat{g}_n})$와 $\boldsymbol{V} = \mathrm{diag}(e^{-\hat{g}_1}, \cdots, e^{-\hat{g}_n})$로 정의된다. 식 (3.2)에서 분자의 h_{tt}를 $\mathrm{tr}(\boldsymbol{H})/n$ 으로 대체하고, 분모의 $h_{tt}e^{\hat{g}_t}$는 $\mathrm{tr}(\boldsymbol{V}^{-1/2}\boldsymbol{H}\boldsymbol{V}^{-1/2})/n$으로 대체하면 다음과 같은 일반화근사교차타당성 함수가 얻어진다.

$$GACV(\boldsymbol{\theta}) = \frac{1}{n}\sum_{t=1}^{n}(z_t e^{-\hat{g}_t} + \hat{g}_t) + \frac{\mathrm{tr}(\boldsymbol{H})}{n}\frac{\sum_{t=1}^{n} e^{-\hat{g}_t}z_t(z_t - e_t^{\hat{g}})}{n - \mathrm{tr}(\boldsymbol{V}^{-1/2}\boldsymbol{H}\boldsymbol{V}^{-1/2})}. \tag{3.3}$$

따라서 최적의 (λ, σ^2)은 식 (3.3)의 GACV 함수를 최소화하는 모수들이다.

4. 모의실험

먼저 모의실험자료를 생성하여 분석해 보고자 한다. 모의실험자료는 오차항의 분포가 정규분포와 t-분포인 자료를 각각 100개씩 생성하여 100회의 모의실험을 시행하였다. 분산추정치 $\hat{\sigma}_t^2$의 초기값은 앞서 언급한 y_t^2의 5일 이동평균을 사용하였으며 모의실험에 사용된 GARCH (1,1) 모형은 $\sigma_t^2 = 0.05 + 0.05y_{t-1}^2 + 0.8\sigma_{t-1}^2$이다. 세 가지 추정법들의 100개의 MSE (mean squared error)의 평균에 대한 결과가 표 4.1과 표 4.2에 주어진다. 그리고 100개의 MAE (mean absolute error)의 평균에 대한 결과가 표 4.3과 표 4.4에 주어진다. 한편 MSE와 MAE는 다음과 같이 정의된다.

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(\hat{\sigma}_t^2 - \sigma_t^2)^2, MAE = \frac{1}{n}\sum_{t=1}^{n}|\hat{\sigma}_t^2 - \sigma_t^2|.$$

표 4.1과 표 4.3의 결과를 살펴보면 정규분포와 t-분포를 따르는 모의실험자료 모두에 대해서 커널기계를 이용한 방법이 SVM를 이용한 방법 보다 MSE값과 MAE값이 더 작은 것을 볼 수 있다. MSE값과 MAE값이 작을수록 자료의 분산과 각 추정법으로 추정한 분산의 차이가 작은 것이므로 커널기계를 이용한 추정법이 SVM를 이용한 추정법 보다 더 좋다고 말 할 수 있다. 표 4.2와 표 4.4의 결과를 살펴보면 자유도가 커질수록 세 가지 추정법에 대한 MSE값과 MAE값이 작아지는 것을 볼 수 있다. 그리고 모든 경우에 커널기계를 이용한 추정법이 가장 작은 MSE값과 MAE값을 제공하는 것을 알 수 있다.

표 **4.1** 정규분포와 t-분포에 대한 GARCH 모형 추정법들의 MSE 비교

	정규분포	t -분포, 자유도 5
최대우도법	**0.0019***	0.8381
SVM	0.0194	0.6969
커널기계	0.0116	**0.2432***

표 **4.2** 자유도가 다른 t-분포에 대한 GARCH 모형 추정법들의 MSE 비교

	자유도 3	자유도 5	자유도 7	자유도 9
최대우도법	196.5908	0.8381	0.0893	0.0582
SVM	465.5489	0.6969	0.2763	0.1134
커널기계*	**51.6427***	**0.2432***	**0.0869***	**0.0455***

표 **4.3** 정규분포와 t-분포에 대한 GARCH 모형 추정법들의 MAE 비교

	정규분포	t -분포, 자유도 5
최대우도법	**0.0347***	0.2843
SVM	0.1122	0.3391
커널기계	0.0740	**0.2785***

표 **4.4** 자유도가 다른 t-분포에 대한 GARCH 모형 추정법들의 MAE 비교

	자유도 3	자유도 5	자유도 7	자유도 9
최대우도법	2.2283	0.2843	0.1667	0.1368
SVM	2.0453	0.3391	0.2449	0.2112
커널기계*	**1.6214***	**0.2785***	**0.1659***	**0.1352***

5. KOSPI 자료분석

이 절에서는 2001년 7월 10일부터 2009년 8월 7일까지 조사된 KOSPI 자료에 대한 GARCH (1,1) 모형의 추정결과를 설명한다. 관측치의 개수는 2000개이다. 그리고 KOSPI 자료는 다음과 같이 로그 수익률로 변환하여 사용된다. 여기서 p_t는 t시점에서 관측된 KOSPI 종가이다.

$$y_t = (\ln p_t - \ln p_{t-1}) \times 100.$$

먼저 분석을 위해 사용된 모형을 살펴보자. 최대우도법을 사용하는 경우에 정규분포를 가정한 GARCH 모형은 $\sigma_t^2 = 0.029 + 0.0798 y_{t-1}^2 + 0.9129\sigma_{t-1}^2$이고 t-분포를 가정한 GARCH 모형은 $\sigma_t^2 = 0.0302 + 0.0792 y_{t-1}^2 + 0.9148\sigma_{t-1}^2$이다. 커널기계를 이용한 경우의 GARCH 모형은 $\sigma^2(\boldsymbol{x}_t) = \exp(\boldsymbol{k}_t\boldsymbol{\alpha} + 0.1088)$이다. 여기서 $\boldsymbol{k}_t$는 커널행렬 $\boldsymbol{K}$ 의 t번째 행벡터이다. GARCH 모형의 추정법들에 대한 MSE값과 R^2값이 표 5.1에 주어진다. MSE와 R^2은 다음과 같이 정의된다.

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(y_t^2 - \hat{y}_t^2)^2, R^2 = 1 - \frac{\sum_{t=1}^{n}(y_t^2 - \hat{y}_t^2)^2}{\sum_{t=1}^{n}(y_t^2 - \frac{1}{n}\sum_{t=1}^{n} y_t^2)^2}.$$

표 5.1을 살펴보면 커널기계의 MSE가 각각 49.5787로서 가장 작고, 커널기계의 R^2값이 0.2182로서 가장 큰 것을 알 수 있다. 그리고 최대우도법을 이용해서 추정한 결과를 보면 GARCH 모형과 EGARCH 모형 모두 t-분포를 가정했을 때의 MSE가 조금 더 작고, R^2값은 조금 더 큰 것을 알 수 있다. 따라서 t-분포를 가정한 자료의 적합도가 더 우수하다고 볼 수 있다. 이 결과는 KOSPI자료의 로그수익률 분포가 정규분포보다는 꼬리가 좀 더 두꺼운 t-분포에 가깝기 때문인 것으로 생각된다.

표 **5.1** 모형추정법들의 MSE와 R^2

	MSE	R^2
최대우도법 - GARCH, 정규분포	57.1407	0.0989
최대우도법 - GARCH, t 분포	57.0191	0.1008
최대우도법 - EGARCH, 정규분포	55.0283	0.1322
최대우도법 - EGARCH, t 분포	54.9386	0.1337
SVM	53.7800	0.1519
커널기계*	**49.5787***	**0.2182***

6. 결론

본 논문의 주요 결과를 요약하면 다음과 같다. 첫째, 모의실험자료에 대한 모형의 추정결과를 고려할 때 정규분포의 경우에 커널기계는 최대우도법 보다 조금 더 큰 MSE값과 MAE값을 제공한다. 그러나 t-분포의 경우에는 다른 두 방법보다 작은 MSE값과 MAE값을 제공한다. 특히 꼬리가 긴 t -분포의 경우에는 다른 두 방법보다 훨씬 작은 MSE값과 MAE값을 제공한다. 둘째, KOSPI의 로그수익률자료에 대한 추정결과를 보면 커널기계가 가장 작은 MSE값을 제공하며, 아울러 커널기계가 가장 큰 R^2값을 제공한다. 따라서 커널기계가 KOSPI 수익률의 변동성을 가장 잘 추정한다고 말할 수 있다. 그리고 최대우도법 보다 SVM이 더 좋은 결과를 보여준다. 본 논문의 결과를 고려할 때 대체로 꼬리가 두꺼운 분포를 갖는 금융시계열자료의 변동성을 추정할 때는 커널기계가 최대우도법과 SVM보다 우수한 방법이 될 것으로 생각된다.

참고문헌

김명직, 장국현 (2002). <금융시계열분석>, 경문사, 서울.

Audrino, F. and Bühlmann, P. (2009). Splines for financial volatility. *Journal of the Royal Statistical Society B*, **71**, 655-670.

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. (2008). Mixed effects kernel binomial regression. *Journal of Korean Data & Information Science Society*, **19**, 1327-1334.

Juutilainen, I. and Röning, J. (2006). *Adaptive modelling of conditional variance function. Proceedings of 17th Symposium of IASC (COMPSTAT 2006)*, Rome, Italy, 1517-1524.

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.

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

Shim, J., Park, H. and Hwang, C. (2009). A kernel machine for estimation of mean and volatility functions. *Journal of Korean Data & Information Science Society*, **20**, 905-912.

Shim, J. and Seok, K. H. (2008). Kernel poisson regression for longitudinal data. *Journal of Korean Data & Information Science Society*, **19**, 1353-1360.

Smola, A. J. and Schoelkopf, B.(1998). *A tutorial on support vector regression. NeuroCOLT2 Technical Report NC-TR-98-030*, Royal Hollow College, University of London, UK.

Xiang, D. and Wahba, G. (1996). A generalized approximate cross validation for smoothing splines with non-Gaussian data. *Statistica Sinica*, **6**, 675-692.

Journal of the Korean Data & Information Science Society
2010, **21**(3), 419–425

한국데이터정보과학회지

Estimating GARCH models using kernel machine learning †

Chang Ha Hwang[1] · Sa Im Shin[2]

[12]Department of Statistics, Dankook University

Received 22 March 2010, revised 8 May 2010, accepted 17 May 2010

Abstract

Kernel machine learning is gaining a lot of popularities in analyzing large or high dimensional nonlinear data. We use this technique to estimate a GARCH model for predicting the conditional volatility of stock market returns. GARCH models are usually estimated using maximum likelihood (ML) procedures, assuming that the data are normally distributed. In this paper, we show that GARCH models can be estimated using kernel machine learning and that kernel machine has a higher predicting ability than ML methods and support vector machine, when estimating volatility of financial time series data with fat tail.

Keywords: GARCH, generalzed approximate cross validation, kernel machine, support vector machine.

† This research was supported by the Graduate Research Assistantship of Dankook University in 2009.
[1] Corresponding author: Professor, Department of Statistics, Dankook University, Yongin 448-701, Korea. E-mail: chwang@dankook.ac.kr
[2] Graduate student, Department of Statistics, Dankook University, Yongin 448-701, Korea.