

서포트벡터기계를 이용한 VaR 모형의 결합

김용태^a, 심주용^b, 이장택^a, 황창하^{1,a}

^a단국대학교 정보통계학과, ^b대구가톨릭대학교 응용통계학과

요약

VaR(Value-at-Risk)는 시장위험을 측정하기 위한 중요한 도구로 사용되고 있다. 그러나 적절한 VaR 모형의 선택에는 논란의 여지가 많다. 본 논문에서는 특정 모형을 선택하여 VaR 예측값을 구하는 대신 대표적으로 많이 사용되는 두개의 VaR 모형인 역사적 모의실험과 GARCH 모형의 예측값들을 서포트벡터기계를 분위수 회귀모형을 이용하여 결합하는 방법을 제안한다.

주요용어: 모형선택, 분위수 회귀, 서포트벡터기계, 역사적 모의실험, GARCH, SVM, Value-at-Risk.

1. 서론

우리나라의 금융기관과 기업은 IMF위기를 겪으면서 엄청난 속도로 금융산업을 개방하고 있다. 이처럼 금융 산업의 개방이 가속화되고 금융의 자율화가 촉진됨에 따라 환율, 추가, 금리 및 상품가격 등 재무변수의 변동성이 더욱 증가하여 시장위험(market risk)이 큰 관심의 대상이 되고 있다. 시장위험관리 기법으로 가장 주목받고 있는 것은 VaR(Value at Risk) 기법으로서 1980년대 후반 주요 금융기관들이 그들의 트레이딩 위험을 계측하기 위하여 처음으로 VaR를 사용하였다. 이후로 VaR의 사용이 폭발적으로 증대되어 현재는 대부분의 파생상품 딜러들이 시장위험을 계측하고 관리하는데 VaR를 사용하고 있다. 자세한 내용은 Jorion (2007)에 설명되어 있다.

VaR는 정상적인 시장조건하에서 일정한 신뢰수준을 가정할 때 일정 기간 동안 보유한 포트폴리오에 대해서 발생할 수 있는 최대손실금액을 의미하며, VaR를 추정하는 방법으로 모수적 방법(Risk Metrics, GARCH 등), 준모수적 방법(Extreme Value Theory, CAViaR 등)과 비모수적 방법(Historical Simulation 등)으로 다양한 모형이 존재하여, 어떤 포트폴리오의 VaR를 추정하는데 각각 다른 모형을 이용하면 서로 다른 추정값을 얻게 된다. 따라서 포트폴리오별로 어떤 모형이 가장 적당한 VaR를 추정해 주는지 선택해야하는 문제에 직면하게 되며, 또 가장 적절하다고 판단되는 모형을 선택하였다고 하더라도 데이터의 변화에 따라서 개별 모형을 이용하여 구한 VaR 예측값은 변화하게 되므로 선택된 모형이 가장 적당한 VaR 모형이라고 말할 수 없을 것이다.

VaR를 예측하기 위해 일반적으로 역사적 모의실험(historical simulation)과 GARCH 모형이 주로 사용되는데 선행연구에 따르면 역사적 모의실험은 VaR를 과대 추정하는 반면에 GARCH 모형은 과소 추정하는 것으로 알려져있다. 본 논문에서는 함수추정에서 우수성이 입증된 서포트벡터기계(support vector machine, SVM)를 이용한 분위수 회귀분석(quantile regression) 기법을 활용하여 두 모형을 결합하는 방법을 제안한다.

이 논문은 2008년도 정부재원(교육인적자원부 학술연구조성사업비)으로 한국학술진흥재단의 지원을 받아 연구되었음(KRF-2008-313-C00159).

¹ 교신저자: (449-791) 경기도 용인시 수지구 죽전동 126 번지 단국대학교 정보통계학과, 교수.
E-mail: chwang@dankook.ac.kr

2. 기존의 VaR 예측모형

VaR는 정상적인 시장(normal market) 여건하에서 주어진 신뢰수준 (confidence level) $1 - \theta$ 로 목표기간(target period) 동안에 보유한 포트폴리오에 대해서 발생할 수 있는 최대손실금액을 의미한다. VaR의 예측값을 구하는 방법은 크게 비모수적 방법과 모수적 방법이 있다. 비모수적 방법은 실제 포트폴리오의 수익률의 분포에 대한 제 100 θ 백분위수를 VaR(θ) 예측값으로 이용한다. 모수적방법은 백분위수 대신 표준편차를 이용하게 된다. 만약 수익률의 분포가 정규분포를 따른다고 가정하면 포트폴리오의 수익률의 평균과 표준편차를 사용하여 다음과 같은 식으로 구하게 된다.

$$\text{VaR}(\theta) = \mu + z_{\theta}\sigma, \quad z_{\theta} = \Phi^{-1}(\theta)$$

위 식에서 보듯이 모수적 방법에서는 표준편차를 추정하는 것이 가장 핵심이 된다.

2.1. 역사적 모의실험

역사적 모의실험을 이용한 VaR 예측방법은 여러 가지가 있다. 본 논문에서는 가장 간단한 역사적 모의실험 방법인 백분위수를 이용해서 VaR를 예측하였다. 즉, $t + 1$ 시점에서의 포트폴리오 수익률을 r_{t+1} 로 나타내자. 역사적 모의실험은 내일의 포트폴리오 수익률 r_{t+1} 의 분포는 과거 m 기간 동안의 포트폴리오 수익률들 $\{r_{t+1-t}\}_{t=1}^m$ 의 경험분포와 같다고 가정한다. 그러면 $\{r_{t+1-t}\}_{t=1}^m$ 의 경험분포의 제 100 θ 백분위수를 $t + 1$ 시점의 VaR의 예측값으로 한다.

2.2. GARCH

Bollerslev (1986)에 의해 제안된 GARCH 모형은 변동성이 큰 기간과 작은 기간이 나뉘어져 있는 시계열의 특성을 명시적으로 반영한 모형이다. 포트폴리오의 수익률을 r_t 라고 할 때 r_t 의 평균방정식(mean equation)과 변동성방정식(volatility equation)은 다음과 같은 GARCH 과정으로 표현된다.

$$r_t = \psi_0 + \sum_{i=1}^p \psi_i r_{t-i} - \sum_{j=1}^q \theta_j a_{t-j} + a_t, \quad a_t = \sigma_t \epsilon_t,$$

$$\sigma_t^2 = \varphi_0 + \sum_{i=1}^u \varphi_i a_{t-i}^2 + \sum_{j=1}^v \theta_j \sigma_{t-j}^2.$$

위의 GARCH 과정은 조건부 평균과 조건부 분산의 1시점앞(one-step-ahead) 예측값 $\hat{r}_t(1), \hat{\sigma}_t^2(1)$ 을 얻기 위해 사용될 수 있다. 이 예측값들을 사용하여 VaR의 예측값을 구할 수 있다. 예컨대 만약 ϵ_t 가 정규분포를 따른다면 $\theta = 0.05$ 일 때 VaR의 1시점앞 예측값은 다음과 같이 구해진다.

$$\text{VaR}_{t+1}(0.05) = \hat{r}_t(1) - 1.65\hat{\sigma}_t(1).$$

만약 ϵ_t 가 자유도 k 의 t -분포를 따른다면 VaR는 다음과 같이 구해진다.

$$\text{VaR}_{t+1}(\theta) = \hat{r}_t(1) - \frac{t_k(\theta)\hat{\sigma}_t(1)}{\sqrt{k/(k-2)}}.$$

3. VaR 예측모형의 결합

금융시장은 일반적으로 매우 복잡하기 때문에 개별 VaR 모형으로 금융시장의 동적변화를 파악하는 것은 일반적으로 어렵다. 금융시장의 복잡성을 파악하기 위해 여러 개의 개별 VaR 모형을 결합하

는 것이 해결책이 될 수 있다. Palit와 Popovic (2000) 및 Ozun과 Cifter (2007)에 따르면 시계열모형을 결합할 때 주로 신경망을 결합알고리즘으로 사용한다. 그러나 본 논문에서 처럼 VaR 예측을 위해 GARCH와 역사적 모의실험을 결합할 때 결합알고리즘으로 신경망을 사용하면 어려운 문제에 봉착할 수도 있다. 왜냐하면 일반적인 신경망은 평균추정을 다루고 오차함수가 미분가능한 대칭함수이지만 VaR 예측을 위한 신경망은 분위수추정을 다루고 오차함수가 다음과 같이 미분불능인 비대칭 점검함수(check function)이기 때문이다.

$$\rho_{\theta}(e_t) = \theta \cdot e_t \cdot I(e_t \geq 0) + (1 - \theta) \cdot e_t \cdot I(e_t < 0), \tag{3.1}$$

여기서 e_t 는 $e_t = r_t - CVaR_t(\theta)$ 로 정의되고 $I(\cdot)$ 는 지시함수(indicator function)이다. 이때 $CVaR_t(\theta)$ 는 t 시점에서 GARCH 모형에 의해 구한 1시점 앞 VaR 예측값을 나타내는 $GVaR_t(\theta)$ 와 역사적 모의실험에 의해 구한 1시점 앞 VaR 예측값을 나타내는 $HVaR_t(\theta)$ 를 결합하여 얻게 되는 VaR를 의미한다.

한편 GARCH와 역사적 모의실험은 서로 다른 형태의 모형이기 때문에 모형결합을 통해 얻어지는 $CVaR_t(\theta)$ 는 두 모형에 의해 구한 VaR들의 비선형 함수일 것이다. 즉, 모형결합을 이용하여 $CVaR_t(\theta)$ 를 구하는 문제는 입력변수가 $GVaR_t(\theta)$ 와 $HVaR_t(\theta)$ 이고 출력변수가 r_{t+1} 일 때의 비선형 분위수 회귀함수를 구하는 문제와 관련되어 있다. 분위수 회귀분석에 대한 자세한 내용은 Koenker와 Bassett (1978)에 설명되어 있다. 본 논문에서는 신경망 대신 최근 분류와 함수추정에서 인기를 얻고 있는 SVM을 이용하여 비선형 분위수 회귀함수를 추정하고 나아가 $CVaR_t(\theta)$ 를 구하고자 한다.

4. SVM 분위수 회귀분석

이 절에서는 Hwang과 Shim (2005)에 소개된 방법과 유사한 QP(quadratic programming)를 이용한 SVM 분위수 회귀함수 추정기법을 설명하고 나아가 IRWLS(iteratively reweighted least squares)를 이용한 SVM 분위수 회귀함수 추정기법을 제안한다. 이 절에서 설명되는 두 방법에서는 Hwang과 Shim (2005)과는 달리 절편항을 독립변수의 회귀계수와 구분하여 다룬다. 왜냐하면 본 논문에서 소개하지는 않지만 선행연구를 통해 Hwang과 Shim (2005)에 소개된 방법은 매우 편익된 결과를 보여주는 것으로 확인되었기 때문이다. SVM에 관한 자세한 내용은 Gunn (1998), Seok 등 (1999), Shim 등 (2009), Smola와 Schölkopf (1998) 및 Vapnik (1998)에 잘 설명되어 있다.

4.1. QP를 이용한 SVM 분위수 회귀함수추정

입력벡터 $\mathbf{x}_i \in R^d$ 와 반응변수 $y_i \in R$ 로 이루어진 학습자료의 집합을 $D = \{\mathbf{x}_i, y_i\}_{i=1}^n$ 라고 정의하자. 특징함수(feature mapping function) $\phi(\cdot) : R^d \rightarrow R^{d_f}$ 는 입력공간(input space)을 고차원 특징공간(feature space)에 대응시킨다. 이때 차원 d_f 는 암묵적으로 정의 되어진다. Mercer (1909)에 따르면 특징공간에서의 내적은 입력공간에서의 커널함수와 동등하다. 즉, $\phi(\mathbf{x}_i)' \phi(\mathbf{x}_j) = K(\mathbf{x}_i, \mathbf{x}_j)$. 커널함수 K 의 종류는 다양하다. 그러나 본 논문에서는 가장 널리 사용되고 있는 가우스 커널을 사용한다.

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{1}{\sigma^2} \|\mathbf{x}_i - \mathbf{x}_j\|^2\right),$$

여기서 σ^2 은 커널모수(kernel parameter)이다.

일반적인 경우의 SVM 분위수 회귀분석을 위해 입력벡터 \mathbf{x} 가 주어졌을 때 반응함수로 비선형 함수 $\mathbf{w}'\phi(\mathbf{x}) + b$ 를 고려한다. 점검함수 $\rho_{\theta}(\cdot)$ 를 사용하여 θ 번째 분위수 회귀함수를 추정하는 최적화 문제는 다음과 같다.

$$\min \frac{1}{2} \mathbf{w}'\mathbf{w} + C \sum_{i=1}^n \rho_{\theta}(y_i - \mathbf{w}'\phi(\mathbf{x}_i) - b), \tag{4.1}$$

여기서 C 는 벌칙모수(penalty parameter)이다. SVM의 기본원리를 사용하면 위의 최적화 문제는 다음과 같은 최적화 문제로 바뀐다.

$$\min \frac{1}{2} \mathbf{w}' \mathbf{w} + C \sum_{i=1}^n (\theta \xi_i + (1 - \theta) \xi_i^*), \quad (4.2)$$

$$\text{제약식} \begin{cases} y_i - \mathbf{w}' \boldsymbol{\phi}(x_i) - b \leq \xi_i, \\ \mathbf{w}' \boldsymbol{\phi}(x_i) + b - y_i \leq \xi_i^*, \\ \xi_i, \xi_i^* \geq 0. \end{cases}$$

위의 최적화 문제를 라그랑주함수를 사용하여 나타내면 다음과 같이 된다.

$$L = \frac{1}{2} \mathbf{w}' \mathbf{w} + C \sum_{i=1}^n (\theta \xi_i + (1 - \theta) \xi_i^*) - \sum_{i=1}^n \alpha_i (\xi_i - y_i + \mathbf{w}' \boldsymbol{\phi}(x_i) + b) - \sum_{i=1}^n \alpha_i^* (\xi_i^* + y_i - \mathbf{w}' \boldsymbol{\phi}(x_i) - b) - \sum_{i=1}^n (\eta_i \xi_i + \eta_i^* \xi_i^*), \quad (4.3)$$

여기서 제약조건 $\alpha_i, \alpha_i^*, \eta_i, \eta_i^* \geq 0$ 이 성립해야 한다. 수식 (4.3)을 변수 $(\mathbf{w}, \xi_i, \xi_i^*)$ 에 대해서 편미분하여 정리하면 위의 최적화 문제는 다시 다음과 같이 바뀐다.

$$\max -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i, x_j) + \sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) \quad (4.4)$$

이때 제약조건은 다음과 같다.

$$\sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0, \quad \alpha_i \in [0, \theta C], \quad \alpha_i^* \in [0, (1 - \theta)C].$$

최적화 문제 (4.4)는 QP를 푸는 문제로 귀결된다.

이제 $\beta_i = \alpha_i - \alpha_i^*$ 를 사용하면 최적화 문제 (4.4)는 다음과 같이 표현된다.

$$\max -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \beta_i \beta_j K(x_i, x_j) + \sum_{i=1}^n \beta_i y_i \quad (4.5)$$

이때 제약조건은 다음과 같다.

$$\sum_i \beta_i = 0, \quad \beta_i \in [(\theta - 1)C, \theta C].$$

최적화 문제 (4.5)를 풀어서 β_i 를 구하면 입력벡터 \mathbf{x} 가 주어질 때 θ 번째 분위수 회귀함수의 추정량은 다음과 같이 된다.

$$q_\theta(\mathbf{x}) = \sum_{i=1}^n \beta_i K(x_i, \mathbf{x}) + b \quad (4.6)$$

이때 절편항 b 는 KKT(Karush-Kuhn-Tucker) 조건을 사용하여 다음과 같이 구해진다.

$$b = \frac{1}{n_s} \sum_{i \in I_s} (y_i - \beta_i K(x_i, \mathbf{x})),$$

여기서 $I_s = \{i : (\theta - 1)C < \beta_i < \theta C\}$ 이고 n_s 는 집합 I_s 의 크기이다.

4.2. IRWLS를 이용한 SVM 분위수 회귀함수 추정

이제 IRWLS를 이용한 SVM 분위수 회귀함수 추정기법을 설명하려고 한다. 이 방법은 QP 기반 방법과는 달리 GCV(generalized cross validation) 기법을 사용하여 최적의 커널모수와 벌칙모수를 구할 수 있고 대용량 자료에도 비교적 쉽게 적용가능하다.

IRWLS 기반 방법은 Nychka 등 (1995)에 의해 0근방의 아주 작은 구간 $(-\delta, \delta)$ 에서 미분가능하도록 수정된 다음과 같은 점검함수 $\rho_{\theta, \delta}$ 를 사용한다.

$$\rho_{\theta, \delta}(e) = \theta \cdot \frac{e^2}{\delta} \cdot I(e \geq 0) + (1 - \theta) \cdot \frac{e^2}{\delta} \cdot I(e < 0). \tag{4.7}$$

따라서 최적화 문제 (4.1)은 β 와 b 에 대해서 다음 수식을 최소화하는 문제로 바뀐다.

$$L(\beta, b) = \frac{1}{2} \beta^t K \beta + C \sum_{i=1}^n \rho_{\theta, \delta}(y_i - K_i \beta - b), \quad \theta \in (0, 1), \tag{4.8}$$

여기서 K_i 는 $n \times n$ 커널행렬 $K = (K(x_i, x_j))_{i,j=1}^n$ 의 i 번째 행을 나타낸다. 수식 (4.8)을 β 와 b 에 관해서 편미분하여 얻은 다음 방정식을 풀면 β 와 b 의 해를 구할 수 있다.

$$\begin{aligned} 0 &= K\beta - CKW y + CKWK\beta + CKW1b \\ 0 &= 1'W y - 1'WK\beta - 1'W1b, \end{aligned} \tag{4.9}$$

여기서 W 는 i 번째 대각원소가

$$w_{ii} = \begin{cases} 2\theta/\delta, & 0 < e_i < \delta, \\ \theta/e_i, & e_i > \delta, \\ 2(1-\theta)/\delta, & -\delta < e_i < 0, \\ (\theta-1)/e_i, & e_i < -\delta \end{cases}$$

인 대각행렬이다. 그리고 $e_i = y_i - K_i \beta - b$ 이다.

선형방정식 (4.9)의 해 β 와 b 는 다음과 같이 초기값 $\beta^{(0)}$ 와 $b^{(0)}$ 를 이용하는 IRWLS 알고리즘을 통해 구해진다.

단계0) β 와 b 의 초기값 $\beta^{(0)}$ 와 $b^{(0)}$ 를 설정한다.

단계1) $\delta, \theta, e_i = y_i - K_i \beta - b$ 를 이용하여 행렬 W 를 구한다.

단계2) 다음 선형방정식을 이용하여 $\beta^{(l)}$ 와 $b^{(l)}$ 을 구한다.

$$\begin{pmatrix} \beta^{(l+1)} \\ b^{(l+1)} \end{pmatrix} = \begin{pmatrix} KWK & \frac{1}{C}K \\ 1'WK & 1'W1 \end{pmatrix}^{-1} \begin{pmatrix} KW \\ 1'W \end{pmatrix} y$$

단계3) 수렴할 때 까지 단계1과 단계2를 반복한다.

이제 SVM 분위수 회귀함수 추정에 결정적으로 영향을 주는 벌칙모수와 커널모수의 값을 결정하는 모형선택 문제를 생각한다. 모형선택을 위해 다음과 같은 CV(cross validation) 함수를 사용한다.

$$CV(\lambda) = \frac{1}{n} \sum_{i=1}^n \rho_{\theta, \delta}(y_i - q_{\theta}^{(-i)}(x_i)),$$

여기서 λ 는 벌칙모수와 커널모수의 집합을 나타내고 $q_{\theta}^{(-i)}(\cdot)$ 는 i 번째 관측치가 제거된 후에 추정된 분위수 회귀함수를 나타낸다. 한편 CV 함수를 계산하는데 시간이 너무 많이 걸리기 때문에 보완책으로 Yuan (2006)이 제안한 GACV(generalized approximate cross validation) 기법을 적용하여 분위수 회귀 함수 추정을 위한 GACV를 구하면 다음과 같다.

$$\text{GACV}(\lambda) = \frac{\sum_{i=1}^n \rho_{\theta, \delta}(y_i - q_{\theta}(\mathbf{x}_i))}{n - \text{trace}(\mathbf{H})}, \quad (4.10)$$

여기서 \mathbf{H} 는 $(q_{\theta}(\mathbf{x}_1), \dots, q_{\theta}(\mathbf{x}_n))' = \mathbf{H}\mathbf{y}$ 를 만족하는 모자행렬(hat matrix)로서 ij 번째 원소는 $h_{ij} = \partial q_{\theta}(\mathbf{x}_i) / \partial y_j$ 이다. 한편 QP를 이용한 SVM 분위수 회귀함수에서는 모자행렬 \mathbf{H} 를 계산할 수 없기 때문에 GACV를 구할 수 없다. 그러나 IRWLS를 이용한 SVM 분위수 회귀함수에서는 위에서 설명한 방법을 따르면 모자행렬을 \mathbf{H} 를 다음과 같이 계산할 수 있기 때문에 GACV를 구할 수 있다.

$$\mathbf{H} = (\mathbf{K}, \mathbf{1}) \begin{pmatrix} \mathbf{K}\mathbf{W}\mathbf{K} & \mathbf{K}/\mathbf{C} \\ \mathbf{1}'\mathbf{W}\mathbf{K} & \mathbf{1}'\mathbf{W}\mathbf{1} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{K}\mathbf{W} \\ \mathbf{1}'\mathbf{W} \end{pmatrix}.$$

5. 실증연구

5.1. 모의실험

SVM 분위수 회귀분석을 이용한 VaR 예측방법과 GARCH 및 역사적 모의실험을 이용한 VaR 예측 방법을 비교하기 위해서 모의실험을 수행한다. 자료는 아래의 GARCH(1,1) 모형을 가정하여 생성한다.

$$\begin{aligned} r_t &= 0.001382 + a_t, \\ \sigma_t^2 &= 0.000003083 + 0.0866a_{t-1}^2 + 0.9032\sigma_{t-1}^2, \\ a_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim t(df.), \quad t = 1, \dots, n, \end{aligned}$$

이때 오차항은 t -분포를 따른다고 가정한다. 자유도(df.)는 3, 5, 10, 100으로 하고 표본의 크기는 $n = 100, 500, 1000$ 으로 정하였다. 모의실험의 과정은 다음과 같다.

- 1) 위의 GARCH(1,1) 모형을 사용해서 $n + 1$ 개의 자료를 생성한다. 이때 처음 n 개의 자료는 VaR를 예측하기 위한 훈련자료이고 $n + 1$ 번째 자료는 비교를 위한 검정자료이다.
- 2) 처음 n 개의 훈련자료에 대해 정규분포 가정하의 GARCH(1,1)과 역사적 모의실험을 이용해서 신뢰 수준 $(1 - \theta)\%$ 하에서 $(\text{GVaR}_{n+1}(\theta), \text{HVaR}_{n+1}(\theta))$ 를 구한다. 여기서 $\theta = 0.05$ 이다.
- 3) 단계 1)과 단계 2)를 100번 반복한다.
- 4) 단계 3)을 통해 얻은 100개의 2-차원 입력벡터 $(\text{GVaR}_{n+1}^{(it)}(\theta), \text{HVaR}_{n+1}^{(it)}(\theta))$ 와 반응변수 $r_{n+1}^{(it)}$ 를 훈련자료로 사용해서 SVM 분위수 회귀함수를 추정한 후 $(\text{GVaR}_{n+1}^{(it)}(\theta), \text{HVaR}_{n+1}^{(it)}(\theta))$ 의 추정값을 결합모형의 예측값 $\text{CVaR}_{n+1}^{(it)}(\theta)$ 로 한다. 여기서 $it = 1, \dots, 100$ 이다.
- 5) 단계 3)과 단계 4)에서 구한 각 VaR 예측값을 이용하여 위반율(violation rate) $1/100 \sum_{it=1}^{100} I(r_{n+1}^{(it)} < \text{GVaR}_{n+1}^{(it)})$, $1/100 \sum_{it=1}^{100} I(r_{n+1}^{(it)} < \text{HVaR}_{n+1}^{(it)})$ 및 $1/100 \sum_{it=1}^{100} I(r_{n+1}^{(it)} < \text{CVaR}_{n+1}^{(it)})$ 를 계산하여 비교한다.

모의실험 결과가 표 1에 주어져있다. 표 1을 살펴보면 역사적 모의실험과 GARCH를 이용한 VaR 예측값 보다 결합모형인 SVM 분위수 회귀모형을 이용한 VaR 예측값이 더 정확하다는 것을 알 수 있

표 1: 모의실험의 VaR 예측값의 위반율

DF		표본 크기			
		100	250	500	1000
3	역사적 모의실험	0.0600	0.0200	0.0300	0.0600
	GARCH	0.0400	0.0200	0.0200	0.0300
	SVM	0.0500	0.0500	0.0400	0.0500
5	역사적 모의실험	0.0400	0.0400	0.0700	0.1000
	GARCH	0.0200	0.0900	0.0400	0.0700
	SVM	0.0400	0.0600	0.0600	0.0700
10	역사적 모의실험	0.0800	0.0400	0.0400	0.0200
	GARCH	0.0600	0.0200	0.0400	0.0300
	SVM	0.0400	0.0500	0.0400	0.0400
100	역사적 모의실험	0.0300	0.0700	0.0700	0.0400
	GARCH	0.0200	0.0700	0.0400	0.0100
	SVM	0.0500	0.0600	0.0500	0.0500

다. 왜냐하면 t -분포의 자유도와 표본의 크기의 대부분의 경우 95% 신뢰수준하에서 SVM 분위수 회귀 분석을 이용한 VaR 예측값의 위반율이 0.05에 가장 가깝기 때문이다.

5.2. 실제자료분석

실제자료의 분석을 위해 한국거래소 홈페이지(<http://www.krx.co.kr/>)에 소개된 2003년 12월 17일부터 2009년 4월 30일까지의 KOSPI200지수의 일별 증가를 사용하는데 전체표본을 3개의 부분표본(subsample)으로 나눈다. 첫 번째 부분표본은 2003년 12월 17일부터 2007년 12월 28일까지 1,000일간의 일별 증가를 포함하고 두 번째 부분표본은 2008년 1월 2일부터 2008년 12월 30일까지 248일간의 일별 증가를 포함한다. 이때 1,000일은 4년간의 거래일에 해당된다. 세 번째 부분표본은 2009년 1월 2일부터 2009년 4월 30일까지 83일간의 일별 증가를 포함한다. VaR의 1시점앞 예측값을 구하고자 하는 기간은 세 번째 부분표본에 해당하는 기간이다. 그림 1은 첫 번째와 두 번째의 부분표본에 대한 로그 수익률의 산점도를 보여주는데 실제로 VaR 예측값을 구하기 위해서 로그 수익률 자료를 사용한다. VaR 예측모형으로 GARCH(1,1)과 역사적 모의실험을 사용하며 결합모형으로 SVM 분위수 회귀모형을 사용한다. 따라서 1시점앞 VaR 예측값들은 다음과 같이 표현된다. 표현의 간결성을 위해 $\theta = 0.05$ 를 생략한다.

$$\{(GVaR_{1248+m}, HVaR_{1248+m}, CVaR_{1248+m})\}_{m=1}^{83}$$

이제 구체적으로 1시점앞 VaR 예측값들을 구하는 과정을 설명하고자 한다. 그 과정을 3단계로 나누어 설명한다.

- 1) 먼저 GARCH(1,1)과 역사적 모의실험을 사용해서 SVM 분위수 회귀함수를 추정하기 위해 필요한 248개의 1시점앞 VaR 예측값들 $\{(GVaR_{1000+i}, HVaR_{1000+i})\}_{i=1}^{248}$ 를 구한다. 예를 들어 첫 번째 수익률 자료부터 1,000번째 수익률 자료까지의 1,000개의 수익률 자료에 대해 제 100% 백분위수를 구하여 $HVaR_{1001}$ 로 하고, 그리고 첫 번째 수익률 자료 r_1 부터 1,000번째 수익률 자료 r_{1000} 까지의 1,000개의 수익률 자료를 이용하여 GARCH(1,1)의 모수들을 추정하여 $GVaR_{1001}$ 을 구한다. $HVaR_{1002}$ 와 $GVaR_{1002}$ 를 구하기 위해서는 1시점을 이동하여 두 번째 수익률 자료 r_2 부터 1,001번째 수익률 자료 r_{1001} 까지의 1,000개의 수익률 자료를 이용하여 같은 방법으로 $GVaR_{1002}$ 와 $HVaR_{1002}$ 를 구한다. 창크기(window size)를 1,000으로 유지하며 계속해서 1시점씩 이동하여 같은 방법으로 $GVaR_{1248}$ 와 $HVaR_{1248}$ 을 구할 때까지 계속한다.

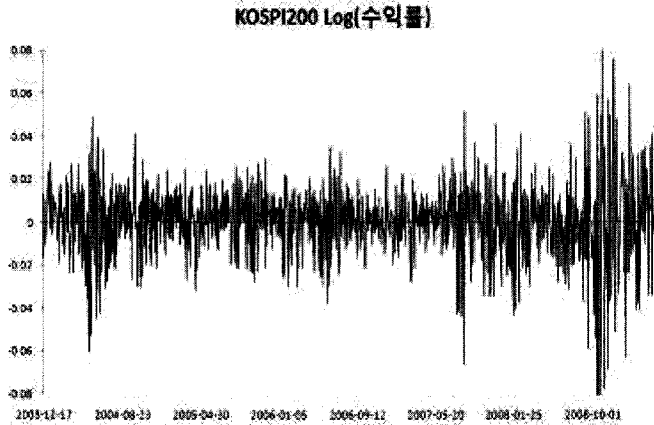


그림 1: 2003년 12월 17일부터 2008년 12월 30일까지의 KOSPI200지수 관련 로그 수익률 자료

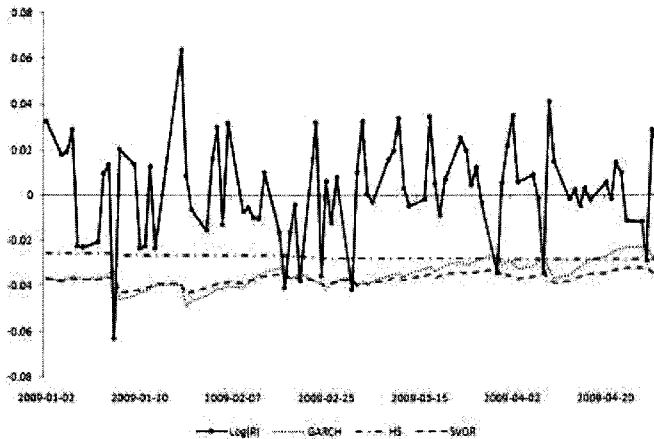


그림 2: 2009년 1월 2일부터 2009년 4월 30일까지 83일간의 KOSPI200지수 관련 로그 수익률 자료 및 VaR 예측값

이렇게 구한 $((GVaR_{1000+i}, HVaR_{1000+i}))_{i=1}^{248}$ 를 입력자료로 수익률 자료 $\{r_{1000+i}\}_{i=1}^{248}$ 를 출력자료로 이용하여 SVM 분위수 회귀함수를 추정한다.

- 2) 249번째 수익률 자료 r_{249} 부터 1,248번째 수익률 자료 r_{1248} 까지의 1,000개의 수익률 자료를 이용하여 $(GVaR_{1249}, HVaR_{1249})$ 를 구하고 이것을 단계 1)에서 추정된 SVM 분위수 회귀함수의 입력으로 적용하여 그 출력값을 $CVaR_{1249}$ 로 한다.
- 3) 단계 1)과 단계 2)의 전 과정을 1시점씩 이동하며 SVM 분위수 회귀모형의 나머지 82개의 VaR 예측값 $CVaR_{1250}, \dots, CVaR_{1331}$ 을 구한다.

그림 2는 세 번째의 부분표본에 대한 로그 수익률의 산점도와 VaR 예측값을 보여준다. 그림 2를 통해서 어떤 방법이 더 좋은지를 판단하기 힘들다. 따라서 VaR 예측모형의 정확성을 검증하는 방법

표 2: 실제자료의 VaR 예측값의 위반율 및 초과량

	예측일수	위반빈도수	위반율	초과량
GARCH	83	7	0.0843	0.0263
HS	83	8	0.0964	0.0446
SVM	83	5	0.0602	0.0142

이 필요하다. 일반적으로 앞에서 설명한 위반율과 초과량(amount of exceedence)을 사용한다. 초과량은 아래에 정의된 VaR Size_t의 평균값이다.

$$\text{VaR Size}_t = \begin{cases} \frac{r_t - \text{VaR}_t}{\text{VaR}_t}, & r_t < \text{VaR}_t, \\ 0, & r_t \geq \text{VaR}_t. \end{cases}$$

표 2는 위반율 및 초과량을 보여준다. 표 2를 살펴볼 때 신뢰수준 95%하에서 결합모형 SVM 분위수 회귀모형을 이용한 VaR 예측값이 개별모형의 VaR 예측값 보다 위반빈도수, 위반율 및 초과량 측면에서 그 값들이 더 작고 더욱이 위반율은 0.05에 더 가깝기 때문에 결합모형 SVM 분위수 회귀모형이 더 정확한 VaR 예측값을 제공함을 알 수 있다.

6. 결론

VaR는 정상적인 시장조건하에서 일정한 신뢰수준을 가정할 때 일정 기간 동안 보유한 포트폴리오에 대해서 발생할 수 있는 최대손실액을 의미하며 시장위험을 측정하는 유용한 통계적 위험관리 기법이다. VaR를 정확하게 예측할 수 있다면 위험에 따른 손실을 최소화 할 수 있을 것이다. 따라서 VaR 예측을 위한 정확한 방법을 유도하는 것은 매우 중요한 문제이다.

본 연구에서는 VaR 예측을 위해 대표적으로 많이 사용되는 역사적 모의실험과 GARCH의 VaR 예측값들을 SVM 분위수 회귀모형을 이용해 결합하는 방법을 제안하고 모의실험과 KOSPI200 지수 관련 실제 수익률 자료에 적용하여 성능을 평가하였다. 위반율과 초과량 측면에서 결합모형 SVM 분위수 회귀모형이 개별 VaR 예측모형 보다 더 정확함을 알 수 있었다.

향후 비대칭 *t*-분포를 사용하는 GARCH와 비대칭 정규혼합 GARCH의 VaR 예측값을 추가하여 신경망 또는 SVM 분위수 회귀모형을 사용하는 두 결합모형의 결과를 비교하는 연구가 필요할 것으로 생각된다.

참고 문헌

- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**, 307–327.
- Gunn, S. (1998). Support vector machines for classification and regression, ISIS Technical Report, University of Southampton.
- Hwang, C. and Shim, J. (2005). A simple quantile regression via support vector machine, *Lecture Notes in Computer Science*, **3610**, 512–520.
- Jorion, P. (2007). *Value at Risk: The New Benchmark for Managing Financial Risk*, McGraw-Hill, New York.
- Koenker, R. and Bassett, G. (1978). Regression quantiles, *Econometrica*, **46**, 33–50.
- Mercer, J. (1909). Functions of positive and negative type, and their connection with the theory of integral equations, *Philosophical Transactions of the Royal Society of London. Series A*, **209**, 415–446.

- Nychka, D., Gray, G., Haaland, P., Martin, D. and O'Connell, M. (1995). A nonparametric regression approach to syringe grading for quality improvement, *Journal of the American Statistical Association*, **90**, 1171–1178.
- Ozun, A. and Cifter, A. (2007). Nonlinear combination of financial forecast with genetic algorithm, MPRA Paper, No.2488.
- Palit, A. K. and Popovic, D. (2000). Nonlinear combination of forecasts using artificial neural network, fuzzy logic and neuro-fuzzy approach, *FUZZ-IEEE*, **2**, 566–571.
- Seok, K., Hwang, C. and Cho, D. (1999). Kernel adatron algorithm for support vector regression, *Communications of the Korean Statistical Society*, **6**, 843–847.
- Shim, J., Hwang, C. and Hong, D. H. (2009). Fuzzy semiparametric support vector regression for seasonal time series analysis, *Communications of the Korean Statistical Society*, **16**, 335–348.
- Smola, A. J. and Schölkopf, B. (1998). A tutorial on support vector regression, NeuroCOLT2 Technical Report, NeuroCOLT.
- Vapnik, V. N. (1998). *Statistical Learning Theory*, Springer.
- Yuan, M. (2006). GACV for quantile smoothing splines, *Computational Statistics & Data Analysis*, **50**, 813–829.

2009년 7월 접수; 2009년 9월 채택

Combination of Value-at-Risk Models with Support Vector Machine

Yongtae Kim^a, Jooyong Shim^b, Jangtaek Lee^a, Changha Hwang^{1,a}

^aDepartment of Statistics, Dankook University

^bDepartment of Applied Statistics, Catholic University of Daegu

Abstract

Value-at-Risk(VaR) has been used as an important tool to measure the market risk. However, the selection of the VaR models is controversial. This paper proposes VaR forecast combinations using support vector machine quantile regression instead of selecting a single model out of historical simulation and GARCH.

Keywords: GARCH, historical simulation, model selection, quantile regression, support vector machine, Value-at-Risk.

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF-2008-313-C00159).

¹ Corresponding author: Professor, Department of Statistics, Dankook University, Jukjeondong, Sujigu, Yongin 449-791, South Korea. E-mail: chwang@dankook.ac.kr