

가중 최소제곱 서포트벡터기계의 혼합모형을 이용한 수익률 기간구조 추정*

노성균¹⁾ 심주용²⁾ 황창하³⁾

요약

수익률 기간구조(term structure of interest rates, 이하 수익률곡선)는 자료의 성격이 경시적(longitudinal)이므로 만기까지 기간과 시간을 동시에 입력변수로 고려해야만 유용하고 효율적인 함수추정이 가능하다. 그러나 이러한 방법은 다루어야 하는 자료가 대용량이기 때문에 대용량 자료에 적합하고 실행속도가 빠른 추정기법을 개발하는 것이 필요하다. 한편 자료에 내재하는 자기상관성 구조 때문에 과대 적합된 추정 결과를 얻기 쉽다. 따라서 본 논문에서는 이러한 문제를 해결하기 위해서 가중 LS-SVM(least squares support vector machine, 최소제곱 서포트벡터기계)의 혼합모형을 제안한다. 미국 재무부 채권에 대한 사례연구를 통해서 추정 결과가 증권시장 붕괴 같은 이례적 사건의 현상을 잘 반영하고 있음을 확인할 수 있었다.

주요용어: 가중 최소제곱 서포트벡터기계, 서포트벡터기계, 수익률 기간구조, 수익률곡선, 혼합모형.

1. 서론

수익률곡선은 수익률과 만기까지 기간의 함수관계로 정의되는데 시장 정보를 잘 나타내는 것으로 알려져 있어서 경기에 대한 기대심리 파악, 수익률에 민감한 자산들의 가격 결정, 다양한 경제변수의 추정 등을 위한 도구로 널리 활용된다. 수익률곡선의 추정은 크게 Vasicek (1977), Cox 등 (1985)의 이론적 접근법과 McCulloch (1971), Nelson과 Siegel (1987), Linton 등 (2000), Krivobokova 등 (2006)의 통계적 접근법이 있다. 이론적 접근법은 수익률 동적모형(dynamics)을 바탕으로 시장에서 무위험차익거래(arbitrage) 이익이 발생하지 않도록 하는 채권가격 모형을 찾는 방법이다. 이러한 모형은 수익률곡선 변화에 대한 일관성 있는 설명이 가능한 장점이 있으나 동적모형의 현실성 혹은 자의성 문제와 함께 실제 시장과의 차이로 무위험차익거래 가능성을 배제하기 어렵다는 문제점이 있다. Ho와

* 본 연구는 한국과학재단 특정기초연구(R01-2006-000-10226-0)지원으로 수행 되었음.

1) (448-160) 경기도 용인시 수지구 죽전동 126번지, 단국대학교 정보컴퓨터과학부, 석사과정.

E-mail: flowflake@naver.com

2) (702-701) 경상북도 경산시 하양읍, 대구가톨릭대학교 응용통계학과, 겸임교수.

E-mail: ds1631@hanmail.net

3) (448-160) 교신저자. 경기도 용인시 수지구 죽전동 126번지, 단국대학교 정보컴퓨터과학부, 교수.

E-mail: chwang@dankook.ac.kr

표 1.1: 표기법 정리

표기	설명	표기	설명
W^{-1}	오차 공분산 행렬	$\varphi(\cdot)$	비선형 사상함수(nonlinear mapping function)
T	전치(transpose)	γ	벌칙 모수(penalty parameter)
w	계수 벡터	σ^2	커널 모수(kernel parameter)
α	라그랑즈 승수 벡터	σ_d^2	상관성이 제거된 오차항의 분산
K	전체 커널행렬	K^{ij}	K 의 i 행, j 열 그룹에 해당하는 부분행렬
λ	모수집합(γ, σ^2)	K^j	K 의 전체행, j 열 그룹에 해당하는 부분행렬
θ	가중치 벡터	θ^j	j 그룹의 가중치
n	전체 자료수	n_i	i 그룹의 자료수
y^i	i 그룹의 출력값	\hat{y}_λ^i	전체자료와 λ 를 이용하여 추정된 i 그룹의 추정치
$\hat{y}_\lambda^{(-i)}$	i 그룹을 뺀 전체자료와 λ 를 이용하여 추정된 i 그룹의 추정치		

Lee (1986) 등은 이러한 점을 고려하여 실제 자료를 입력으로 활용하는 모형을 제안하였으나 Backus 등 (1996)에 의해 여전히 무위험차익거래 가능성을 배제하기 어렵다는 것이 지적된 바 있다. 통계적 접근법은 스플라인과 같은 유연한 함수로 자료를 가장 잘 적합하도록 하는 할인함수 혹은 현물수익률 함수를 찾도록 하는 방법이다. 이러한 모형은 실제 자료의 대부분이 설명 가능하고 비모수 통계 기법의 경우 모형에 대한 사전 정의를 하지 않기 때문에 잘못된 모형설정에 의한 가격오차나 헤징위험을 피할 수 있다는 장점이 있다. 그러나 이론적 접근법과는 반대로 수익률곡선 변화에 대한 일관성 있는 설명이 쉽지 않은 단점이 있다. 이와 같이 수익률곡선의 추정법은 상호 보완적인 장단점을 갖으며 때로는 통계적 접근법이 이론적 모형에 대한 타당성 연구의 바탕이 되기도 한다.

수익률 자료는 시장에서 거래되는 채권들에 대한 경시적 조사의 형태로 얻어지기 때문에 수익률곡선의 추정에 해당하는 횡단면분석과 장단기 채권의 수익률 추이를 살피는 시계열분석이 모두 가능하다. 그러나 단순히 횡단면분석 혹은 시계열분석처럼 한 가지 정보만을 활용하는 방법 보다는 주어진 정보를 모두 활용하는 방법이 더 유용하고 효율적인 추정을 가능하게 하여 선호된다. 또한 잘 평활화된 결과는 이상점 발생 및 수익률곡선의 평행이동과 기울기 변화 같은 현상을 한눈에 파악할 수 있으므로 스프레드 거래같은 투자전략에 유용하게 사용 되어질 수 있다. 그러나 이렇게 만기까지 기간과 시간이 입력변수인 수익률 함수를 평활화 하여 추정하는 방법은 다루어야 하는 자료가 대용량이기 때문에 특별한 추정방법이 요구된다. 또한 개별 채권의 시간에 따른 수익률 변화를 살펴보면 자기상관성이 존재하는 것을 알 수 있는데 이러한 구조를 무시하는 경우 과대 적합된 결과를 얻게 되는 문제점이 있다. 따라서 본 논문에서는 자기상관성에 대한 해법으로 가중 LS-SVM을 사용하고 대용량자료에 대한 해법으로 혼합모형을 사용하는 가중 LS-SVM의 혼합모형을 제안한다. 그리고 이러한 방법의 유용성을 실증적인 측면에서 알아보기 위해서 미국 재무부 채권에 대한 사례연구를 실시하였다.

본 논문은 다음과 같이 구성된다. 우선 논문에 사용되는 표기법을 표 1.1에 정리한다.

2절에서는 가중 LS-SVM에 대해서 간략히 알아보고, 3절에서는 가중 LS-SVM의 혼합모형과 개선된 교차타당성 함수를 제안한다. 4절에서는 가중 LS-SVM의 혼합모형으로 미국 재무부 채권의 수익률곡선을 추정하고 증권시장 붕괴와 같은 이례적 사건과 연관지어 해석해본다. 마지막으로 5절에서는 논문의 결론을 맺는다.

2. 가중 LS-SVM

p 개의 입력변수와 1개의 출력변수에 대한 자료집합이 $Z = \{\mathbf{x}_l, y_l\}_{l=1}^n$, $\mathbf{x}_l \in \mathbb{R}^{p+1}$, $y_l \in \mathbb{R}$ 로 주어졌을 때 상수항을 포함하는 가중 LS-SVM 모형은 $\mathbf{Q} = (\varphi(\mathbf{x}_1), \dots, \varphi(\mathbf{x}_n))$, $\mathbf{y} = (y_1, \dots, y_n)^T$ 로 정의하여 다음과 같이 표현할 수 있다.

$$\mathbf{y} = \mathbf{Q}^T \mathbf{w}. \quad (2.1)$$

오차항 \mathbf{e} 가 평균이 $\mathbf{0}$ 이고 분산이 $\sigma_d^2 \mathbf{W}^{-1}$ 인 다변량 분포를 따른다고 가정하면 가중 LS-SVM의 목적함수 $J(\mathbf{w}, \mathbf{e})$ 는 다음과 같고,

$$\min_{\mathbf{w}, \mathbf{e}} J(\mathbf{w}, \mathbf{e}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{\gamma}{2} \mathbf{e}^T \mathbf{W} \mathbf{e}, \quad (\text{제약식 : } \mathbf{y} = \mathbf{Q}^T \mathbf{w} + \mathbf{e}) \quad (2.2)$$

라그랑주 함수 $L(\mathbf{w}, \mathbf{e}; \boldsymbol{\alpha})$ 는 다음과 같이 구성할 수 있다.

$$\begin{aligned} L(\mathbf{w}, \mathbf{e}; \boldsymbol{\alpha}) &= J(\mathbf{w}, \mathbf{e}) - \boldsymbol{\alpha}^T (\mathbf{Q}^T \mathbf{w} + \mathbf{e} - \mathbf{y}) \\ &= \frac{1}{2} \mathbf{w}^T \mathbf{w} + \frac{\gamma}{2} \mathbf{e}^T \mathbf{W} \mathbf{e} - \boldsymbol{\alpha}^T (\mathbf{Q}^T \mathbf{w} + \mathbf{e} - \mathbf{y}). \end{aligned} \quad (2.3)$$

최적조건은 식 (2.3)을 $\mathbf{w}, \mathbf{e}, \boldsymbol{\alpha}$ 에 대해서 편미분한 다음의 결과로부터 주어지고,

$$\begin{cases} \frac{\partial L}{\partial \mathbf{w}} = \mathbf{0} \Rightarrow \mathbf{w} = \mathbf{Q} \boldsymbol{\alpha}, \\ \frac{\partial L}{\partial \mathbf{e}} = \mathbf{0} \Rightarrow \mathbf{e} = \frac{1}{\gamma} \mathbf{W}^{-1} \boldsymbol{\alpha}, \\ \frac{\partial L}{\partial \boldsymbol{\alpha}} = \mathbf{0} \Rightarrow \mathbf{Q}^T \mathbf{w} + \mathbf{e} = \mathbf{y} \end{cases} \quad (2.4)$$

Mercer (1909)의 조건을 만족하는 커널함수($K(\mathbf{x}_i, \mathbf{x}_j) = \varphi(\mathbf{x}_i)^T \varphi(\mathbf{x}_j)$)를 사용해서 $\mathbf{K} = \mathbf{Q}^T \mathbf{Q}$ 로 정의하여 정리하면 다음의 결과를 얻는다.

$$\boldsymbol{\alpha} = \left(\mathbf{K} + \frac{1}{\gamma} \mathbf{W}^{-1} \right)^{-1} \mathbf{y}. \quad (2.5)$$

위 결과를 식 (2.4)와 식 (2.1)에 대입하여 정리하면 $\hat{\mathbf{y}}$ 은 최종적으로 다음의 식으로 구해진다는 것을 알 수 있다.

$$\begin{aligned} \hat{\mathbf{y}} &= \mathbf{K} \boldsymbol{\alpha} \\ &= \mathbf{K} \left(\mathbf{K} + \frac{1}{\gamma} \mathbf{W}^{-1} \right)^{-1} \mathbf{y}. \end{aligned} \quad (2.6)$$

참고로 본 논문에서는 여러 종류의 커널함수 중에서 가우스 커널 ($K_{ij} = \varphi(\mathbf{x}_i)^T \varphi(\mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2/\sigma^2)$, $i, j = 1, \dots, n$)의 사용을 전제로 한다.

3. 가중 LS-SVM의 혼합모형과 교차타당성 함수

자료의 크기에 따라서 커널행렬의 차원도 달라지기 때문에 일반적으로 대용량 자료인 경우 커널행렬과 관련된 역행렬의 직접적인 연산이 어려운 경우가 많다. 따라서 이와 같은 경우 Suykens 등 (1999)의 CG(conjugate gradient), Keerthi와 Shevade (2003)의 SMO(sequential minimal optimization) 등과 같은 특별한 알고리즘이 필요하게 된다. 본 논문에서는 이러한 문제의 해법으로 혼합모형을 이용한다. 이 방법은 자료의 상관성을 고려해서 전체자료를 M 개 그룹으로 나누고 각 그룹의 가중 LS-SVM에 가중치를 주어 혼합하는 방법이다.

3.1. 가중 LS-SVM의 혼합모형

j 그룹의 자료집합 \mathbf{Z}^j 로부터 구한 라그랑주 승수 벡터가 $\boldsymbol{\alpha}^j$ 일 때 $\mathbf{k}^j = (K(\mathbf{x}, \mathbf{x}_1^j), \dots, K(\mathbf{x}, \mathbf{x}_{n_j}^j))$ 로 정의하고 식 (2.6)을 이용하여 정리하면 혼합모형에 의한 \hat{y} 은 다음과 같이 표현할 수 있다.

$$\hat{y}(\mathbf{x}) = \sum_{j=1}^M \theta^j \mathbf{k}^j \boldsymbol{\alpha}^j. \quad (3.1)$$

$\boldsymbol{\alpha}^j$ 는 $n_j \ll n$ 이므로 식 (2.5)를 이용하면 쉽게 구할 수 있고 최적의 $\boldsymbol{\theta}$ 는 다음의 목적함수로부터 구할 수 있다.

$$D = \frac{1}{n} \sum_{i=1}^M \left\| \mathbf{y}^i - \sum_{j=1}^M \theta^j \mathbf{K}^{ij} \boldsymbol{\alpha}^j \right\|^2, \quad (\text{제약식 : } \theta^j \geq 0, j = 1, \dots, M). \quad (3.2)$$

위 식을 최소화하는 양수 벡터 $\boldsymbol{\theta}$ 를 구하는 것은 쉽지 않으므로 D 가 이차식임을 감안하여 \mathbf{H} 를 원소 $H_{ij} = \boldsymbol{\alpha}^{iT} \mathbf{K}^{-iT} \mathbf{K}^{-j} \boldsymbol{\alpha}^j$ 인 $M \times M$ 행렬, \mathbf{f} 를 원소 $f_j = \mathbf{y}^T \mathbf{K}^{-j} \boldsymbol{\alpha}^j$ 인 $M \times 1$ 행렬로 정의해서 정리한 다음의 식에 QP(quadratic programming)를 적용해서 수치적으로 구하게 된다.

$$\min_{\boldsymbol{\theta}} \left(\frac{1}{2} \boldsymbol{\theta}^T \mathbf{H} \boldsymbol{\theta} - \mathbf{f}^T \boldsymbol{\theta} \right), \quad (\text{제약식 : } \boldsymbol{\theta} \geq 0). \quad (3.3)$$

이때 다수의 가중치(θ^j)가 0이 되어 실제로 훈련후 검정에 쓰이는 자료의 수를 줄여주는 효과가 있기 때문에 빠른 추정이 가능한 장점이 있다. 마지막으로 알고리즘을 정리하면 다음과 같다.

가중 LS-SVM의 혼합모형

1. 전체 자료를 상관성을 고려해서 M 개의 그룹으로 나눈다.

2. 각 그룹의 자료집합 Z^j 로부터 라그랑즈 승수 벡터 α^j 를 구한다.
3. 식 (3.3)을 QP로 풀어서 최적의 θ 를 구한다.
4. 식 (3.1)을 이용해서 최적의 \hat{y} 을 구한다.

3.2. 개선된 교차타당성 함수

모형선택시 일반적인 leave-one-out CV(cross validation, 교차타당성)함수는 비효율적이므로 가중 LS-SVM혼합모형에 맞추어 소요시간을 효과적으로 줄일 수 있는 개선된 CV함수를 고려해 보자. 모수집합 λ 에 대한 일반적인 leave-one-out CV함수는 다음과 같다.

$$CV(\lambda) = \frac{1}{n} \sum_{i=1}^M \|y^i - \hat{y}_\lambda^{(-i)}\|^2. \quad (3.4)$$

위 식은 모든 λ 후보에 대해 $\hat{y}_\lambda^{(-i)}$, $i = 1, \dots, M$ 의 연산이 요구되므로 모형선택시 많은 시간이 소요된다. CV함수의 재구성을 위해서 \hat{y}_λ 와 $\hat{y}_\lambda^{(-i)}$ 를 정리하면 다음과 같다.

$$\hat{y}_\lambda^i = \sum_{j=1}^M \theta^j K^{ij} \alpha^j = \sum_{j=1}^M \theta^j K^{ij} \left(K^{jj} + \frac{1}{\gamma} (W^j)^{-1} \right)^{-1} y^j, \quad (3.5)$$

$$\hat{y}_\lambda^{(-i)} = \sum_{j=1, j \neq i}^M \theta^j K^{ij} \alpha^j = \sum_{j=1, j \neq i}^M \theta^j K^{ij} \left(K^{jj} + \frac{1}{\gamma} (W^j)^{-1} \right)^{-1} y^j. \quad (3.6)$$

식 (3.5)와 식 (3.6)을 비교하면 $\hat{y}_\lambda^{(-i)}$ 이 다음과 같이 다시 표현되는 것을 알 수 있다.

$$\begin{aligned} \hat{y}_\lambda^{(-i)} &= \hat{y}_\lambda^i - \theta^i K^{ii} \left(K^{ii} + \frac{1}{\gamma} (W^i)^{-1} \right)^{-1} y^i \\ &= \hat{y}_\lambda^i - A_\lambda^i y^i. \end{aligned} \quad (3.7)$$

따라서 식 (3.4)는 다음과 같이 다시 쓸 수 있다.

$$CV(\lambda) = \frac{1}{n} \sum_{i=1}^M \|(\mathbf{I}_{n_j} + A_\lambda^i) y^i - \hat{y}_\lambda^i\|^2. \quad (3.8)$$

위 식은 모든 λ 후보에 대해 \hat{y}_λ^i 의 연산이 1번씩만 요구되기 때문에 모형선택 시간을 효과적으로 줄일 수 있다.

4. 미국 재무부 채권에 대한 사례분석

본 논문에서는 미국 재무부 STRIPS(separate trading of registered interest and principal of securities)자료에 대해 사례분석을 하였다. STRIPS는 이표채를 원금과 이자로 나누어 유통시키는 방법으로 채권의 종류에 따라 달라지는 이표효과(coupon effect)를 분리해내므로

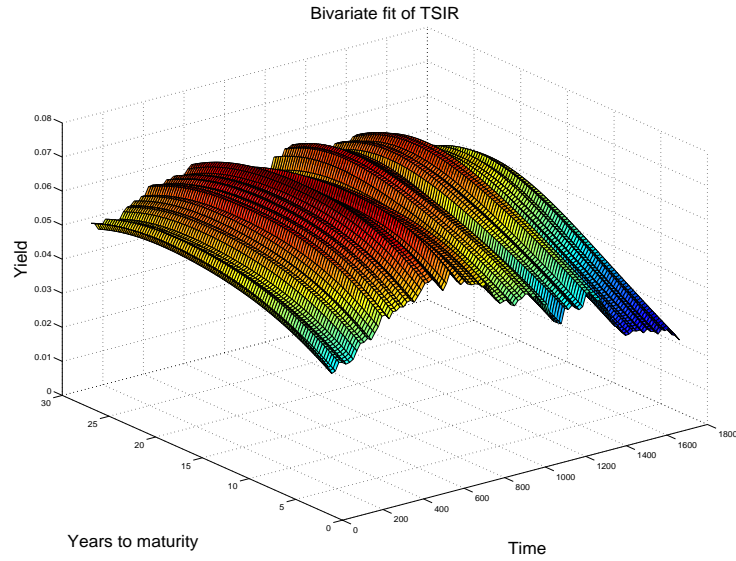


그림 4.1: 추정된 시간과 만기까지 기간 그리고 수익률의 함수(1998년 7월 ~ 2003년 7월)

현물수익률 함수를 바로 추정하는 본 논문의 방법에 적합하다. 자료는 총 126,251개로 만기가 2003년 8월부터 2033년 5월까지인 107개의 채권을 대상으로 1998년 7월부터 2003년 7월까지 조사하였다. 시장에서 t 시점에 만기까지 기간이 m 인 제로쿠폰채(zero coupon bond)가 $P_{t,m}$ 의 가격으로 거래된다면 연속복리에 의한 수익률(continuously compounded yield)은 $y = -\log(P_{t,m})/m$ 로 쉽게 계산된다. 시간과 만기까지 기간 그리고 수익률사이에는 어떠한 함수 관계가 존재하므로 이것을 $g(\cdot)$ 로 가정하면 다음과 같이 표현할 수 있다.

$$y = g(t, m) + \epsilon.$$

가중 LS-SVM 혼합모형으로 $g(t, m)$ 를 추정하기 위해서는 우선 전체 자료의 상관성을 파악해야 한다. 개별 채권의 시간에 따른 수익률 변화를 살펴보면 자기상관성이 존재하는 것을 알 수 있으나 채권들 사이에는 어떠한 상관성이 존재한다고 보기 어렵다. 따라서 전체 자료를 채권별로 그룹짓고 시간에 따라 정렬하면 자기상관성에 대한 고려가 가능하다. 사전 조사 결과 총 107개 그룹은 평균적으로 $\rho = 0.85$ 인 AR(1)모형을 따르는 것으로 나타났으므로 j 번째 그룹의 $n_j \times n_j$ 행렬 $(W^j)^{-1}$ 은 다음과 같이 표현할 수 있다.

$$(W^j)^{-1} = \frac{1}{1 - \rho^2} \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n_j-1} \\ \rho & 1 & \rho & \dots & \rho^{n_j-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n_j-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n_j-1} & \rho^{n_j-2} & \rho^{n_j-3} & \dots & 1 \end{bmatrix}, \quad \rho = 0.85. \quad (4.1)$$

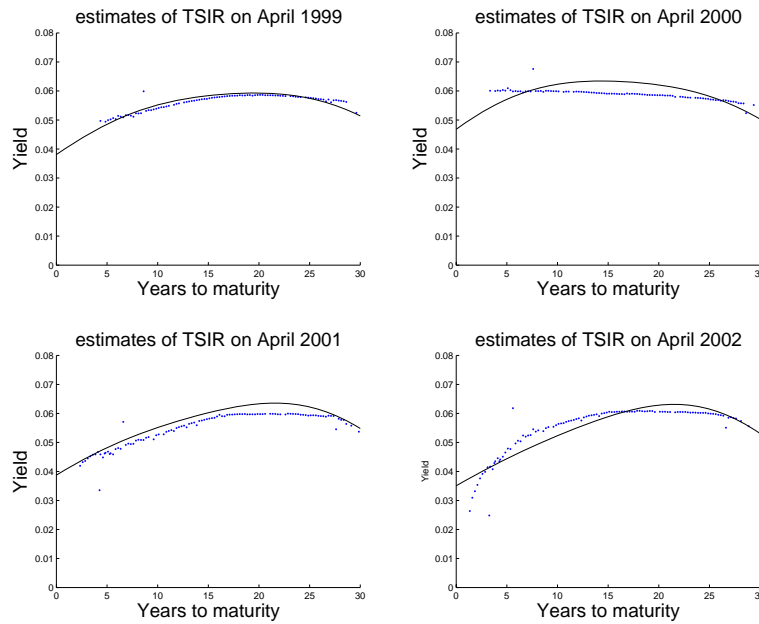


그림 4.2: 2000년 3월 미국 증시 붕괴 전후의 추정된 수익률곡선 변화(실제 자료값(점), 추정된 수익률곡선(선))

미국의 나스닥지수는 2000년 3월(600일 근처에 해당) 5,132.52포인트에서 2002년 10월 1,108.49포인트로 74.8% 가까이 떨어졌다. 통산 세번째 증권시장 붕괴로 총 8조 달러가 사라졌고 채권시장에도 적지 않은 영향을 주었으리라 짐작할 수 있다. 모형선택에 소요되는 시간을 줄이기 위해서 각 그룹에서 200개씩의 자료를 임의추출 하였고 총 21,400개의 자료에 대해 개선된 CV함수를 적용하여 최적의 모수집합($\gamma = 100, \sigma^2 = 250$)을 얻었다. 이 모수집합으로 시간과 만기까지 기간 그리고 수익률의 함수를 추정하면 그림 4.1과 같다. 추정된 결과는 잘 평활화 되었으므로 수익률 곡선 자체의 시간에 따른 변화를 한 눈에 파악할 수 있다. 증권시장의 영향으로 600일 근처에서 단기 채권의 수익률이 장기 채권의 수익률을 웃도는 특이현상이 발생한 것을 알 수 있다. 이러한 상호적인 현상들을 조금 더 자세히 파악하기 위해서 그림 4.1을 시간축과 만기까지 기간축으로 잘라서 살펴보았다. 그림 4.2는 1999년부터 2002년까지 매년 4월의 수익률곡선을 추정한 것이다. 1999년 4월의 수익률곡선은 만기까지 기간이 커질 수록 수익률도 커지는 일반적인 형태이다. 그러나 증시 붕괴 한 달 후인 2000년 4월에는 단기채권의 수익률이 장기채권의 수익률을 웃도는 현상을 보여준다. 그 후 2001년, 2002년 4월의 수익률곡선은 다시 일반적인 우상향 형태를 회복하고 있음을 알 수 있다. 이렇듯 가중 LS-SVM 혼합모형으로 추정된 수익률곡선의 결과는 증권시장 붕괴와 같은 사건의 전형적인 모습을 잘 나타내고 있다. 그림 4.3은 단기물과 장기물의

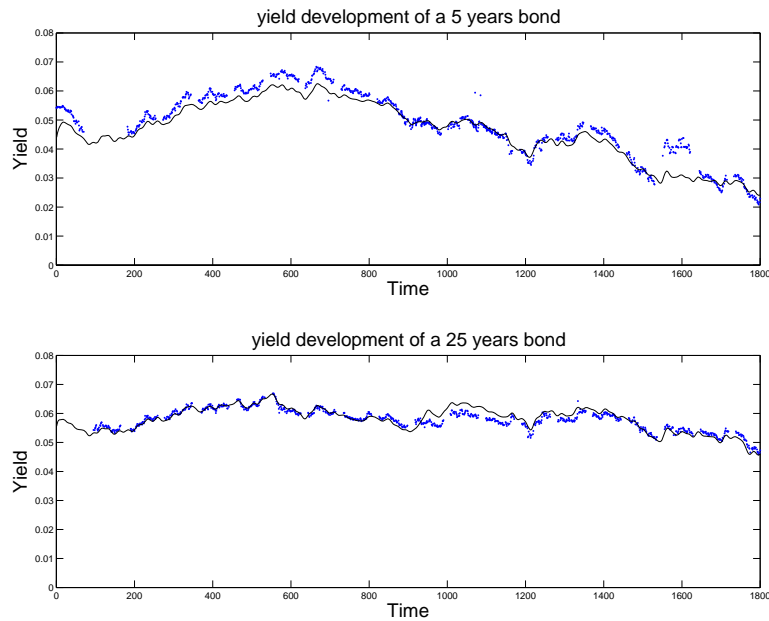


그림 4.3: 단기물(5년, 상)과 장기물(25년, 하)의 시간에 따른 수익률 변화(근사 자료값(점), 시간에 따른 추정 수익률의 변화(선))

시간에 따른 수익률 변화를 추정한 것이다. 12만 여개의 자료들 중에서도 만기까지 기간이 동일한 자료의 수는 극히 적기 때문에 추정치와의 직접적인 비교는 불가능하다. 그러나 수익률곡선을 살펴보면 한 달 이내의 수익률 변화가 미미하여 만기까지 기간 ± 36.5 일의 자료가 같은 값을 갖는다고 가정한다면 실제 자료의 근사치로 사용할 수 있다. 단기물의 경우 2000년 3월 붕괴직전에는 수익률이 0.02올랐다가 그 이후 0.04나 크게 떨어져 변동폭이 아주 큰 모습을 보여주고 있다. 그러나 장기물의 경우 이와는 반대로 0.02정도의 작은 변동만 있어 시장상황에 민감하지 않다는 것을 알 수 있다. 이와 같이 단기물과 장기물은 그 특성이 다르기 때문에 시장상황에 따라 수익률곡선이 여러 가지 형태를 띠게 되는 것을 알 수 있다. 이와 관련해서 순수기대이론, 유동성 선호이론, 시장분할이론 등 여러 가설이 있지만 의견이 분분하고 본 논문의 목적과는 맞지 않으므로 이를 이용한 해석 및 검증은 다루지 않는다.

5. 결론

본 논문에서는 수익률 자료의 특성을 고려하여 수익률 곡선을 추정하는 방법으로 가중 LS-SVM 혼합모형과 개선된 CV함수를 새롭게 제시하였다. 미국 재무부 채권에 대한 사례

분석은 만족할만한 결과를 보여 주었으며 이러한 방법은 다음과 같은 장점을 갖는다. 첫째, 자료기간 내에는 동일한 모수집합을 유지하기 때문에 수익률곡선을 추정할 때 마다 모수가 달라지게 되는 현상을 막을 수 있다. 둘째, 평활화된 결과를 얻음으로서 이상점 발생 및 수익률곡선의 평행이동과 기울기 변화 같은 현상을 한 눈에 파악할 수 있다. 셋째, 만기까지 기간이 동일한 자료가 적음에도 불구하고 순수하게 해당년물의 시간에 따른 수익률 변화의 추정이 가능하다. 넷째, 모형에 대한 사전정의를 하지 않기 때문에 잘못된 모형설정에 의한 가격오차나 헤징위험을 피할 수 있다. 통계적 접근법은 이론적 접근법과 상호보완적인 장단점을 갖으며 때로는 이론적 모형에 대한 타당성 연구의 바탕으로 사용되기도 한다는 점을 고려할 때 다양한 장점을 갖는 본 논문의 의의를 찾을 수 있다.

참고문헌

- Backus, D., Foresi, S. and Zin, S. (1996). Arbitrage opportunities in arbitrage-free models of bond pricing, *Journal of Business & Economic Statistics*, **16**, 13–26.
- Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985). A theory of the term structure of interest rates, *Econometrica*, **53**, 385–407.
- Ho, T. S. Y. and Lee, S. B. (1986). Term structure movements and pricing interest rate contingent claims, *The Journal of Finance*, **41**, 1011–1029.
- Keerthi, S. S. and Shevade, S. K. (2003). SMO algorithm for least squares SVM formulations, *Neural Computation*, **15**, 487–507.
- Krivobokova, T., Kauermann, G. and Archontakis, T. (2006). Estimating the term structure of interest rates using penalized splines, *Statistical Papers*, **47**, 443–459.
- Linton, O., Mammen, E., Nielsen, J. and Tanggaard, C. (2000). Estimating yield curves by kernel smoothing methods, *Journal of Econometrics*, **105**, 185–223.
- McCulloch, J. H. (1971). Measuring the term structure of interest rates, *The Journal of Business*, **44**, 19–31.
- 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*, Ser. A, **209**, 415–446.
- Nelson, C. R. and Siegel, A. F. (1987). Parsimonious modelling of yield curves, *Journal of Business*, **60**, 473–489.
- Suykens J. A. K., Lukas, L., Van Dooren, P., De Moor, B. and Vandewalle, J. (1999). Least squares support vector machine classifiers: A large scale algorithm, European Conference on Circuit Theory and Design(ECCTD' 99), 839-842.
- Vasicek, O. (1977). An equilibrium characterization of the term structure, *Journal of Financial Economics*, **5**, 177–188.

[2007년 10월 접수, 2007년 12월 채택]

Estimating the Term Structure of Interest Rates Using Mixture of Weighted Least Squares Support Vector Machines

Sungkyun Nau¹⁾ Jooyong Shim²⁾ Changha Hwang³⁾

ABSTRACT

Since the term structure of interest rates (TSIR) has longitudinal data, we should consider as input variables both time left to maturity and time simultaneously to get a more useful and more efficient function estimation. However, since the resulting data set becomes very large, we need to develop a fast and reliable estimation method for large data set. Furthermore, it tends to overestimate TSIR because data are correlated. To solve these problems we propose a mixture of weighted least squares support vector machines. We recognize that the estimate is well smoothed and well explains effects of the third stock market crash in USA through applying the proposed method to the US Treasury bonds data.

Keywords: Mixture, support vector machine, term structure of interest rates, weighted least squares support vector machine, yield curve.

1) Graduate Student, Division of Information and Computer Science, Dankook University, 126 Jukjeon-dong, Suji-gu, Yongin-si, Gyeonggi-do 448-160, Korea.

E-mail: flowflake@naver.com

2) Professor, Dept. of Applied Statistics, Catholic University of Daegu, Hayang-eup, Gyeongsan-si 702-701, Korea.

E-mail: ds1631@hanmail.net

3) Corresponding author. Professor, Division of Information and Computer Science, Dankook University, 126 Jukjeon-dong, Suji-gu, Yongin-si, Gyeonggi-do 448-160, Korea.

E-mail: chwang@dankook.ac.kr