

# Multiple-threshold asymmetric volatility models for financial time series

Hyo Ryoung Lee<sup>a</sup>, Sun Young Hwang<sup>1,a</sup>

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

---

## Abstract

This article is concerned with asymmetric volatility models for financial time series. A generalization of standard single-threshold volatility model is discussed via multiple-threshold in which we specialize to two-threshold case for ease of presentation. An empirical illustration is made by analyzing S&P500 data from NYSE (New York Stock Exchange). For comparison measures between competing models, parametric bootstrap method is used to generate forecast distributions from which summary statistics of CP (Coverage Probability) and PE (Prediction Error) are obtained. It is demonstrated that our suggestion is useful in the field of asymmetric volatility analysis.

Keywords: asymmetric volatility, multiple-threshold, parametric bootstrap

---

## 1. 서론

주식, 환율, 금리를 포함하는 금융 시계열의 수익률(return)은 시점에 따라 변동성(volatility, 조건부 분산)이 달라지는 특징을 갖는다. 금융 시계열에서 변동성은 위험(risk)의 측도이므로 리스크 관리, 포트폴리오 선택, 파생상품의 가치 평가 등의 부문에서 중요하게 다루어진다 (Choi 등 2020, 2021; Yoon과 Hwang, 2015). 일반적인 변동성의 특징으로는 변동성이 높고 낮은 기간이 군집을 이루면서 나타나는 변동성 (기간별) 집중 현상(volatility cluster)과 자산 가격의 상승과 하락에 따라 변동성이 다르게 반응하는 비대칭 효과(asymmetric effect, leverage effect)가 있다 (Tsay, 2010). Engle (1982)은 변동성을 과거의 관측값 제곱들로 회귀분석 모형화하는 자기회귀 조건부 이분산(auto regressive conditional heteroskedasticity; ARCH) 모형을 제안하였다. Bollerslev (1986)는 일반화된 조건부 이분산(GARCH) 모형 점화식으로 ARCH 모형을 일반화하였다.

금융 시계열 변동성을 나타내는 점화식으로 많이 이용하는 수식으로 ARCH(1) 모형이 있다. 시계열  $\{\epsilon_t\}$ 가 다음과 같은 점화식으로 축차적으로(sequentially) 생성될 때 ARCH(1) 모형으로 정의한다.

$$\epsilon_t = \sqrt{h_t}e_t, \quad h_t = \alpha_0 + \alpha_1\epsilon_{t-1}^2, \quad (1.1)$$

여기서  $h_t = \text{Var}(\epsilon_t | F_{t-1})$ 는 시점  $t$ 에서의 변동성이며 오차  $\{e_t\}$ 는 평균 0, 분산 1인 iid과정이다. 관련 모수는 모두 비음(non-negative)을 가정하며 모형 (1.1)은 변동성이  $\{\epsilon_t\}$ 에 대해 대칭을 보여주고 있다. 최근 금융 시계열 분야에서는 실증적 자료 분석을 통해 변동성이 비대칭이라는 점이 널리 인정되고 있으며 이에 맞추어 비대칭 효과를 반영할 수 있는 비대칭 변동성 모형의 연구가 활발하게 진행되고 있다(stylized fact, Tsay, 2010, Ch. 3, Choi 등, 2021).

---

This work was supported by a grant from the National Research Foundation of Korea (NRF-2021R1F1A1047952).

<sup>1</sup> Corresponding author: Department of Statistics, Sookmyung Women's University, Cheongpa-ro 47-gil 100, Yongsan-gu, Seoul 04310, Republic of Korea. E-mail: shwang@sookmyung.ac.kr

변동성 모형은 임계점(threshold)을 통해 비대칭화(asymmetrization) 할 수 있으며 기존의 임계점을 이용한 변동성(volatility) 모형은 하나의 임계점을 이용하여 특정 고정된 지점에서 비대칭성이 이루어졌다. 본 연구에서는 기존 임계점 관련 모형을 다중 임계점(multiple-threshold)을 가진 모형으로 확장하여 금융 시계열의 비대칭 변동성 모형을 개선하고자 한다. 특히, 논의의 간편성을 위해서 두 개의 임계점(double-threshold)을 가진 비대칭 변동성 모형을 집중적으로 연구하고자 한다. 모형 간의 비교 지표로 Log Likelihood, AIC, BIC를 사용하였고 모수적-부스트랩(parametric bootstrap)방법을 통해 예측분포를 구성하고 이를 통해 CP(coverage probability)와 예측오차(prediction error) 요약 통계량을 만들어서 예측력을 비교 평가한다. 미국 금융 시계열 자료인 S&P500 자료를 이용해서 제안된 모형의 예제 분석을 실시하였다.

## 2. 비대칭 변동성 임계점 모형

본 절에서는 표준적인 TARCH(1) 모형과 다중 임계점을 가진 비대칭 ARCH(1) 점화식을 소개하고자 한다. 비대칭 변동성 임계점 모형에 대해서는 Tsay (2010), Hwang 등 (2010), Choi 등 (2020) 연구와 문헌을 바탕으로 기본적인 개념과 수식을 정리하였다. 논의의 간편성 및 모형의 응용성을 고려하여 다중 임계점을 가진 비대칭 ARCH(1) 모형은 두 개의 임계점 모형을 중심으로 서술하고자 하였다.

### • 비대칭 변동성 TARCH (Threshold ARCH) 모형식 : TARCH(1)

Rabemananjara와 Zakoian (1993)이 제안한 threshold ARCH (T-ARCH) 모형식은 다음과 같다.

$$\epsilon_t = \sqrt{h_t}e_t, \quad h_t = \alpha_0 + \alpha_{11}(\epsilon_{t-1}^+)^2 + \alpha_{12}(\epsilon_{t-1}^-)^2, \quad (2.1)$$

여기서 임계점 함수는  $\epsilon^+ = \max(\epsilon, 0)$ 과  $\epsilon^- = \max(-\epsilon, 0)$ 이며 계수  $\alpha_{11}, \alpha_{12}$ 는 비대칭 특성을 반영하는 모수이다.  $\alpha_{12} > \alpha_{11}$ 인 경우, 수익률 영(zero)을 중심으로, 같은 양이면 양의 수익률을 가질 때보다 음의 수익률을 가질 때가 변동성이 더 큰(즉, 금융자산의 위험이 더 큰) 비대칭 효과(leverage effect)를 의미한다.  $\alpha_{11} = \alpha_{12}$ 이면 변동성이 수익률의 부호에 영향을 받지 않는 표준적인 ARCH 모형이다 (Choi 등, 2021).

### • 다중 임계점을 가진 비대칭 변동성 ARCH 모형식

식(2.1)의 일반화로서 다음과 같은 다중 임계점 모형을 고려할 수 있다.

$$h_t = \sum_{i=1}^k \alpha_{0i} I_{R_i} + \sum_{i=1}^k \alpha_{1i} \epsilon_{t-1}^2 I_{R_i},$$

여기서  $R_1, \dots, R_k$ 는 실수공간  $(-\infty, \infty)$ 의 분할(exclusive partition)이며  $I$ 는 지시함수(indicator function)이다. 특별한 경우로서  $k = 2$ 이고  $R_1 = (-\infty, 0), R_2 = [0, \infty)$ 인 경우는 식 (2.1)의 표준적인 TARCH(1) 모형이다. 다중 임계점 모형을 실제 자료에 적용하기 위해서는 분할공간인  $R_1, \dots, R_k$ 를 선택해야 하는 문제가 있다. 본 연구에서는 금융 시계열의 변동성에 적용하기 위해서 수익률 영(zero)과 인접한 값  $\delta$ 와  $-\delta$ 에서 임계점이 있는 다음과 같은 모형을 고려한다.

$$\epsilon_t = \sqrt{h_t}e_t, \quad h_t = [\alpha_{01} + \alpha_{11}\epsilon_{t-1}^2] I_{[\epsilon_{t-1} < -\delta]} + [\alpha_{02} + \alpha_{12}\epsilon_{t-1}^2] I_{[|\epsilon_{t-1}| < \delta]} + [\alpha_{03} + \alpha_{13}\epsilon_{t-1}^2] I_{[\epsilon_{t-1} > \delta]}. \quad (2.2)$$

즉, 세분화한 비대칭 변동성 효과를 반영하기 위해 두 개의 임계점( $\delta$ 와  $-\delta$ )을 고려하여 나누어지는 세 개의 수익률 구간에 대해 각각 기울기가 다른 변동성 모형을 구축하였다. 본 연구에서는 식 (2.2)를 분석하고자 하며 이 식을 Asymmetric ARCH(1), 즉, “**A-ARCH(1)**”으로 부르기로 한다. A-ARCH(1) 모형 수식의 특징은 세 개의 이차함수가 구간별로 결합한 형태이며 임계점  $\delta$ 와  $-\delta$ 에서 변동성 함수( $h_t$ )가 불연속(discontinuous)이며 따라서 미분 불가능(non-differentiable)하므로 모수 추정에 어려움이 생긴다.

### • 모수 추정 방법

모형식 (2.2)의 두 개의 임계점( $\delta$ 와  $-\delta$ )을 가진 비대칭 ARCH(1) 모형을 추정하기 위하여 Nelder-Mead(NM), BFGS, CG, SANN의 4가지 방법을 사용하였다. 아래에서 소개한 모수 추정 방법은 R을 통해서 구현가능하다.

- (1) **NM**방법은 Nelder와 Mead가 제안한 방법으로 다차원 독립변수들을 이용하여 목적함수 최적화에 사용하는 방법이며 목적함수의 최솟값 혹은 최댓값을 찾기 위한 수치적인 방법으로 도함수를 알 수 없는 비선형 최적화 문제 상황에서 사용된다. 1965년 Nelder와 Mead가 처음으로 소개한 방법으로 아메바 방법이라고도 부른다. 항상 수렴이 되는 강력한 방법이나 상대적으로 반복 횟수가 많아 계산 속도가 느리다는 단점이 있다 (Nelder와 Mead, 1965).
- (2) **BFGS**는 Quasi-Newton 방법으로 제한적인 컴퓨터 메모리를 사용한다. 비선형 최적화 문제 해결을 위한 반복적인 방법으로 한정된 경계 내에서 모수 학습을 하는 경우 사용하기 적절하다. 반복을 시행할 때마다 함수의 Hessian 근사를 통해 Newton 방법에 근사한다. 1970년 “Broyden-Fletcher-Goldfarb-Shanno”가 제안한 알고리즘 방법이다 (Guerrout 등, 2018; Byrd 등, 1995).
- (3) **CG**는 켈레-기울기법(Conjugate Gradient)으로 n차원의 최소화 문제를 (n-1)차원의 초평면에서 해를 찾는 문제로 차원을 하나씩 줄여가는 반복을 실행한다. 주어진 시간 내에 정확한 값을 구하기 어렵다면 최대한 가까운 값을 추정한다. 따라서 편미분 방정식이나 최적화 문제를 수치적으로 풀 때 자주 사용하는 방법이다. 1964년 Fletcher와 Reeves가 제안하였다. 일반적으로 BFGS 방법보다 취약하나 행렬을 저장하지 않으므로 훨씬 큰 최적화 문제에 유용하다 (Fletcher와 Reeves, 1964).
- (4) **SANN**은 1992년 Belisle이 제안한 simulated-annealing의 변형된 방법이다. 미분 불가능한 함수에 대해서도 가능하며 지정된 최대 반복 횟수에 대해 항상 함숫값을 사용하고 평가한다 (Belisle, 1992; Nocedal와 Wright, 1999).

### 예제 : S&P 500 변동성 분석

S&P 500(Standard & Poor’s 500 index)는 500개의 대형기업 주식을 포함한 지수이다. 세계 3대 신용평가 기관으로 불리고 있는 미국의 스탠더드 앤드 푸어스(Standard and Poors, S&P)가 작성한 주가지수이다. 다우 존스 지수와 마찬가지로 뉴욕증권거래소에 상장되어있으며 포함되는 종목 수가 많아 가장 많이 활용되는 대표적인 주가지수 중 하나이다. 식 (2.2) 시계열 모형의 적합을 위해 사용된 자료는 2000년 1월 3일에서 2020년 12월 30일까지 총 20년간 5283개의 일별 증가 데이터  $\{P_t\}$ 이며  $\{r_t\}$ 는 일별 수익률(단위 : %), 즉,

$$r_t = (\ln P_t - \ln P_{t-1}) \times 100$$

이다. 이 기간 평균 수익률  $\mu = 0.01783$ 을 고려하여 다음 모형식을 고려하였다.

$$r_t = 0.01783 + \epsilon_t,$$

여기서  $\epsilon_t$ 는 식 (2.2)를 따른다.

R의 optim 함수를 사용하여 위에서 서술한 네 가지 방법을 A-ARCH 모형에 적합하였다. 모수추정을 위한 초기값으로 세 구간 모두에서 변동성 함수의 절편( $\alpha_0$ )은 0.2, 이차함수 기울기( $\alpha_1$ )은 0.9로 하였다. 수익률의 임계점  $\delta$ 에 대해서는  $\delta$ 를 0.1(%)에서 5.0(%)까지 0.1(%) 간격으로 변경하면서 총 50개의 시나리오에 대한 다중 임계점을 가진 비대칭 ARCH 모형 모수를 추정하였다. 모형 비교 기준인 AIC와 BIC를 비교하였을 때, 가장 성능이 우수한 모형은  $\delta = 0.3$ 일 때 CG method를 사용한 다중 임계점을 가진 비대칭 ARCH(1) 모형이었다(AIC는 3.10819, BIC는 3.11566으로 가장 작은 값을 갖는다). 두 번째로 성능이 좋은 모형은  $\delta = 0.3$ 일 때 Nelder-Mead method를 사용한 다중 임계점을 가진 비대칭 ARCH(1) 모형이며 AIC는 3.10821, BIC는 3.11567

Table 1: Estimated equations of A-ARCH(1)

Model	Estimated equation
A-ARCH(1) NM ( $\delta = 0.3$ )	$r_t = 0.017830 + \epsilon_t, \epsilon_t = \sqrt{h_t} e_t$ $h_t = [1.276735 + 0.44604\epsilon_{t-1}^2]I_{[\epsilon_{t-1} < -\delta]} + [0.841437 + 0.683264\epsilon_{t-1}^2]I_{[ \epsilon_{t-1}  < \delta]} + [1.048002 + 0.195544\epsilon_{t-1}^2]I_{[\epsilon_{t-1} > \delta]}$
A-ARCH(1) NM ( $\delta = 0.5$ )	$r_t = 0.017830 + \epsilon_t, \epsilon_t = \sqrt{h_t} e_t$ $h_t = [1.477069 + 0.378016\epsilon_{t-1}^2]I_{[\epsilon_{t-1} < -\delta]} + [0.908972 + 0.383447\epsilon_{t-1}^2]I_{[ \epsilon_{t-1}  < \delta]} + [1.019274 + 0.201246\epsilon_{t-1}^2]I_{[\epsilon_{t-1} > \delta]}$
A-ARCH(1) NM ( $\delta = 1.0$ )	$r_t = 0.017830 + \epsilon_t, \epsilon_t = \sqrt{h_t} e_t$ $h_t = [1.850974 + 0.3369224\epsilon_{t-1}^2]I_{[\epsilon_{t-1} < -\delta]} + [0.919580 + 0.620390\epsilon_{t-1}^2]I_{[ \epsilon_{t-1}  < \delta]} + [1.087851 + 0.189866\epsilon_{t-1}^2]I_{[\epsilon_{t-1} > \delta]}$
A-ARCH(1) BFGS ( $\delta = 1.0$ )	$r_t = 0.017830 + \epsilon_t, \epsilon_t = \sqrt{h_t} e_t$ $h_t = [1.869461 + 0.326991\epsilon_{t-1}^2]I_{[\epsilon_{t-1} < -\delta]} + [0.909358 + 0.662606\epsilon_{t-1}^2]I_{[ \epsilon_{t-1}  < \delta]} + [1.078936 + 0.187950\epsilon_{t-1}^2]I_{[\epsilon_{t-1} > \delta]}$
A-ARCH(1) CG ( $\delta = 0.3$ )	$r_t = 0.017830 + \epsilon_t, \epsilon_t = \sqrt{h_t} e_t$ $h_t = [1.277281 + 0.446394\epsilon_{t-1}^2]I_{[\epsilon_{t-1} < -\delta]} + [0.834809 + 0.944040\epsilon_{t-1}^2]I_{[ \epsilon_{t-1}  < \delta]} + [1.047897 + 0.195567\epsilon_{t-1}^2]I_{[\epsilon_{t-1} > \delta]}$
A-ARCH(1) CG ( $\delta = 1.0$ )	$r_t = 0.017830 + \epsilon_t, \epsilon_t = \sqrt{h_t} e_t$ $h_t = [1.869439 + 0.326994\epsilon_{t-1}^2]I_{[\epsilon_{t-1} < -\delta]} + [0.909358 + 0.662605\epsilon_{t-1}^2]I_{[ \epsilon_{t-1}  < \delta]} + [1.078919 + 0.187953\epsilon_{t-1}^2]I_{[\epsilon_{t-1} > \delta]}$
A-ARCH(1) SANN ( $\delta = 0.3$ )	$r_t = 0.017830 + \epsilon_t, \epsilon_t = \sqrt{h_t} e_t$ $h_t = [1.241682 + 0.452776\epsilon_{t-1}^2]I_{[\epsilon_{t-1} < -\delta]} + [0.827341 + 1.780048\epsilon_{t-1}^2]I_{[ \epsilon_{t-1}  < \delta]} + [1.050751 + 0.214599\epsilon_{t-1}^2]I_{[\epsilon_{t-1} > \delta]}$
A-ARCH(1) SANN ( $\delta = 1.0$ )	$r_t = 0.017830 + \epsilon_t, \epsilon_t = \sqrt{h_t} e_t$ $h_t = [1.889081 + 0.318944\epsilon_{t-1}^2]I_{[\epsilon_{t-1} < -\delta]} + [0.905099 + 0.646722\epsilon_{t-1}^2]I_{[ \epsilon_{t-1}  < \delta]} + [1.047685 + 0.182995\epsilon_{t-1}^2]I_{[\epsilon_{t-1} > \delta]}$

로 나타났다. 다양한 A-ARCH(1) 모형 중, AIC와 BIC를 기준으로 성능이 뛰어난(가장 작은 값을 갖는) 상위 8개 모형 식을 Table 1에 정리하였다.

News Impact Curve (NIC)는  $t-2$ 시점 이하의 로그 수익률은 모두 상수로 간주하고 가로축을  $\epsilon_{t-1}$ , 세로축을 변동성  $h_t$ 로 나타낸 2차원 그래프이다. NIC는 비대칭 효과(leverage effect) 여부를 시각적으로 확인할 수 있는 도구이다. NIC에 대한 좀 더 자세한 내용은 Yoon 등 (2014)을 참고하기 바란다. A-ARCH(1)의 네 개의 추정 방법에 따른 NIC 그림을 Figure 1에 제시하였다. 네 가지 NIC 그림에서 A-ARCH(1)의 추정 모수는 다르나 그림의 형태는 비슷하다.  $\delta$ 와  $-\delta$ 에서 불연속이 발생하고 있고 각각의 구간에서 기울기가 다르다.  $\epsilon_{t-1} < \delta$  부분에서의 기울기가  $\epsilon_{t-1} > \delta$ 부분에서의 기울기보다 크므로 비대칭 효과(leverage effect)가 존재한다는 것을 시각적으로 확인할 수 있다.

### 3. 추정 모형의 예측력 평가 : 모수적 붓스트랩

경쟁 모형(competing models)의 비교 평가는 시계열 예측 분야에서 매우 중요한 연구 이슈이다. 모형 선택에 널리 이용되는 AIC(BIC)는 예측력을 평가하는 통계량은 아니며 시계열 예측력 평가 통계량 얻기 위해서 Miguel과 Olave (1999)가 제안한 모수적 붓스트랩(Parametric Bootstrap, 이하 **PB**로 표기)을 사용하여 미래 시점 변동성의 예측분포를 구성한다. 모수적 붓스트랩에 대해서는 Choi 등 (2021)의 내용을 참고하여 서술하였으며 자세한 내용은 Kim과 Hwang (2020)과 Tsay (2010, 4장)을 참고하기 바란다. PB는 추정된 모수를

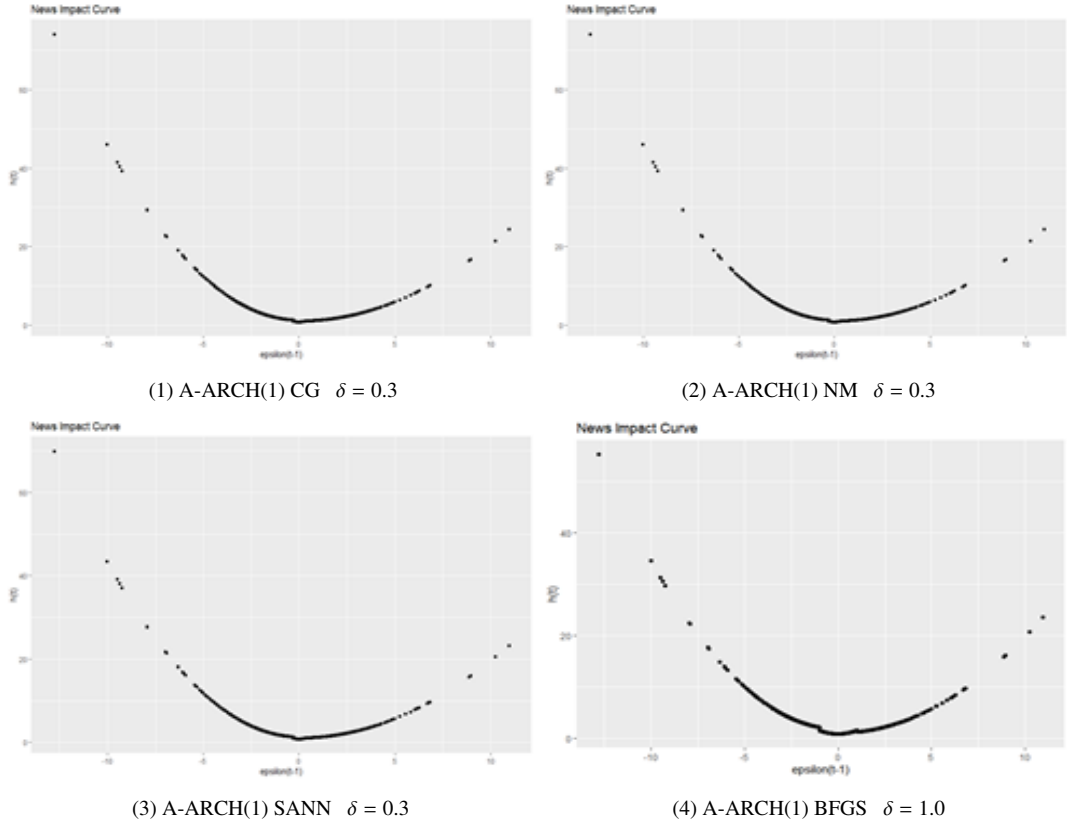


Figure 1: NIC of A-ARCH(1)

참값으로 가정하고 고정하여 분석하게 되므로 모수 추정 과정에서의 불확실성은 고려하지 않는다. 다중 임계점을 가진 A-ARCH(1) 모형 (2.2)에 대한 PB 알고리즘은 다음과 같다. 모형 점화식은

$$r_t = \mu + \epsilon_t, \quad \epsilon_t = \sqrt{h_t}e_t$$

$$h_t = [\alpha_{01} + \alpha_{11}\epsilon_{t-1}^2]I_{[\epsilon_{t-1} < -\delta]} + [\alpha_{02} + \alpha_{12}\epsilon_{t-1}^2]I_{[|\epsilon_{t-1}| < \delta]} + [\alpha_{03} + \alpha_{13}\epsilon_{t-1}^2]I_{[\epsilon_{t-1} > \delta]}$$

이며 자료는  $T$ 개로 구성된  $\{r_1, r_2, \dots, r_T\}$ 이다.

• 모수적 붓스트랩 알고리즘

- (1) 합리적인 방법을 이용하여 모수를 추정한다. 여기서는 2절에서 소개한 Nelder-Mead, BFGS, CG, SANN의 4가지 방법을 사용하였으며 추정식은 Table 1에 수록하였다.
- (2) 추정된 모형을 통해 표준화 잔차(standardized residuals)  $\hat{\epsilon}_t = \epsilon_t / \sqrt{h_t}$ 을 계산한다.
- (3) 중심화 잔차  $\hat{\epsilon}_t - \hat{\epsilon}_\bullet$ 를 생성한 후, 중심화 잔차의 누적분포함수  $\hat{F}$ 를 구한다.

$$\hat{F}(x) = \frac{1}{T} \sum_{t=1}^T I_{[\hat{\epsilon}_t - \hat{\epsilon}_\bullet, \infty)}(x),$$

여기서  $\hat{\epsilon}_\bullet = \sum_{t=1}^T \hat{\epsilon}_t / T$ 이다.

Table 2: Coverage Probabilities for 8 A-ARCH(1) models

Model	95%			99%		
	$B = 300$	$B = 1000$	$B = 3000$	$B = 300$	$B = 1000$	$B = 3000$
TARCH(1)	90.7%	91.0%	91.3%	<b>96.3%</b>	96.3%	96.3%
A-ARCH(1) NM ( $\delta = 0.3$ )	<b>91.7%</b>	92.0%	92.0%	95.7%	<b>96.7%</b>	96.7%
A-ARCH(1) NM ( $\delta = 0.5$ )	<b>91.7%</b>	91.7%	<b>92.3%</b>	96.0%	96.3%	96.7%
A-ARCH(1) NM ( $\delta = 1.0$ )	91.3%	91.7%	92.0%	96.0%	96.0%	96.3%
A-ARCH(1) BFGS ( $\delta = 1.0$ )	91.0%	91.0%	91.3%	96.0%	96.3%	96.3%
A-ARCH(1) CG ( $\delta = 0.3$ )	<b>91.7%</b>	<b>92.3%</b>	<b>92.3%</b>	95.3%	96.3%	<b>97.3%</b>
A-ARCH(1) CG ( $\delta = 1.0$ )	90.7%	<b>92.3%</b>	<b>92.3%</b>	<b>96.3%</b>	96.3%	96.7%
A-ARCH(1) SANN ( $\delta = 0.3$ )	90.0%	91.7%	91.7%	96.0%	96.3%	96.7%
A-ARCH(1) SANN ( $\delta = 1.0$ )	89.7%	91.0%	91.0%	95.0%	96.3%	96.7%

- (4) 중심화 잔차 분포  $\hat{F}$  으로 부터 랜덤 추출하여 미래값  $e_{T+j}^*$ ,  $j = 1, 2, \dots, l$  을 생성한다. 여기서  $j = 1$  은 1-시차 후 예측이며  $j = l$  은  $l$ -시차 후 예측을 위한 잔차이다. 본 분석에서는  $l = 1$  인 경우만 고려한다.
- (5) 추정된 모형식과 랜덤 추출한 잔차를 기반으로 새로운 시계열 표본  $\{r_{T+j}^*\}$ 과 (변동성) 조건부 분산  $\{h_{T+j}^*\}$  을 생성한다.
- (6)  $e_{T+j}^*$ ,  $j = 1, 2, \dots, l$ ,  $\{r_{T+j}^*\}$ ,  $\{h_{T+j}^*\}$  생성을  $B$ 번 반복하여 예측 분포(predictive distribution)를 구한다. 본 연구에서는  $B$ 를 300, 1000, 3000으로 선택하였다.

#### • PB 알고리즘을 이용한 예측구간 평가

모수적 붓스트랩을 통해서 Table 1에 수록한 8개의 A-ARCH(1) 경쟁 모형에 대한 예측력을 비교 평가해보도록 하자. 1-시차 후 예측( $l = 1$ )을 분석하였으며 반복 수는  $B = 300$ ,  $B = 1000$ ,  $B = 3000$ 을 사용하였다. 최근  $n$  개의 데이터를 고려하여  $T - 1, T - 2, \dots, T - n - 1$  시점에서  $B$ 번의 모수적 붓스트랩을 계산하여 새로운 로그 수익률 데이터를 생성하고 1-시차 후 예측분포를 구성한다. 이 예측 분포로부터 (명목 신뢰도 95% 및 99%) 예측구간을 구하고 이 구간이 실제 관측 수익률의 포함 여부를 확인한다. 최근  $n$ 개의 수익률의 포함 비율인 Coverage Probability(CP)를 계산한다. 전체 로그 수익률 5282개 데이터에서 최근 300개를 검증 표본으로 사용하여 평가에 이용하였으며 결과를 Table 2에 정리하였다. 참고로, R의 rugarch 패키지를 이용해서 구한 식 (2.1)의 TARCH(1) 추정식은 다음과 같다.

$$r_t = 0.015333 + \epsilon_t, \quad \epsilon_t = \sqrt{h_t} e_t$$

$$h_t = 0.890520 + 0.651470(\epsilon_{t-1}^-)^2 + 0.360165(\epsilon_{t-1}^+)^2.$$

TARCH(1) 추정식의 AIC = 3.11390이고 BIC = 3.11890로서 Table 1에 수록한 상위 8개 모형의 AIC와 BIC 모형보다 큰 값이었으며 TARCH(1) 모형의 예측력 평가를 위한 모수적 붓스트랩 결과도 Table 2와 Table 3에 A-ARCH(1)과 함께 수록하였다.

명목 신뢰도에 가까운 CP를 제공하는 모형이 우수한 모형이다. 95% 예측구간의 경우,  $\delta = 0.3$ 인 CG A-ARCH(1) 모형이 명목 신뢰도 95%에 가장 가까웠으며 99%의 예측구간의 경우, 명목 신뢰도 99%에 가까운 모형이 반복수  $B$ 에 따라 다른 결과가 나타났으나  $B = 3000$ 일 때  $\delta = 0.3$ 인 CG A-ARCH(1) 모형이 가장 우수하였다. 따라서 예측구간의 포함 비율인 CP-기준에서  $\delta = 0.3$ 인 CG A-ARCH(1) 모형이 우수하다는 결론을 내릴 수 있다. TAR(1) 모형은 명목 신뢰도 99%의 반복수  $B = 300$ 에서 우수한 CP를 제공하고 있다.

#### • PB 알고리즘을 이용한 예측오차(prediction error; PE) 분석

모수적 붓스트랩(PB)을 통해 얻은  $\{r_{t+1}\}$ 의 반복(replication)  $B$ 개의 표본평균을  $r_{t+1}$ 의 예측값으로 선택하여 예측오차를 구한다. 최근  $n = 300$ 일로부터 계산한 예측오차로부터 MSE(Mean Squared Error), MAD(Mean Absolute Deviation, 평균절대오차), MAPE(Mean Absolute Percentage Error, 평균절대백분위오차) 및 WMAPE(Weighted Mean Absolute Percentage Error, 가중평균절대백분위오차)를 계산한다. WMAPE는 분모가 영(zero)에 가까운 값들이 있는 경우에 왜곡될 수 있는 MAPE값의 대응으로 이용된다. 예측 오차(300개)를 이용한 요약 통계량은 Table 3에 정리하였다. 사용한 요약통계량의 계산 공식은 다음과 같다.

$$\begin{aligned} \text{MSE} &= \frac{1}{n} \sum_{i=0}^{n-1} (X_{T+1+i} - X_{T+i}(1))^2, \\ \text{MAD} &= \frac{1}{n} \sum_{i=0}^{n-1} |X_{T+1+i} - X_{T+i}(1)|, \\ \text{MAPE} &= \frac{1}{n} \sum_{i=0}^{n-1} \left| \frac{X_{T+i}(1)}{X_{T+1+i}} - 1 \right| \times 100, \\ \text{WMAPE} &= \frac{\sum_{i=0}^{n-1} (|X_{T+i}(1) - X_{T+1+i}|)}{\sum_{i=0}^{n-1} (|X_{T+1+i}|)} \times 100. \end{aligned}$$

Table 3에 의하면  $B = 300$ 에서 MSE/MAD기준으로  $\delta = 0.3$ 인 CG를 사용한 A-ARCH(1) 모형,  $\delta = 1.0$ 인 CG를 사용한 A-ARCH(1) 모형을 최적으로 선택하고 있다. 기준 MAPE/WMAPE에 의하면  $\delta = 1.0$ 인 Nelder-Mead를 사용한 A-ARCH(1) 모형,  $\delta = 0.3$ 인 CG를 사용한 A-ARCH(1) 모형을 최적으로 선택하고 있다.  $B = 1000$ 에서는 MSE/MAD기준으로  $\delta = 0.3$ 인 CG A-ARCH(1) 모형,  $\delta = 1.0$ 인 SANN A-ARCH(1) 모형을 최적으로 선택하고 있다. MAPE/WMAPE기준으로는  $\delta = 1.0$ 인 BFGS A-ARCH(1) 모형,  $\delta = 1.0$ 인 SANN A-ARCH(1) 모형을 최적으로 선택하고 있다.  $B = 3000$ 에서는 MSE/MAD기준으로  $\delta = 1.0$ 인 CG A-ARCH(1) 모형이 가장 우수하다. MAPE/WMAPE기준으로는  $\delta = 0.5$ 인 Nelder-Mead A-ARCH(1) 모형,  $\delta = 1.0$ 인 CG A-ARCH(1) 모형을 최적으로 선택하고 있다. 전반적으로는 임계점  $\delta = 1.0$ 을 우수한 모형으로 판단할 수 있다. TAR(1) 모형은 모든 반복수  $B$ 에서 상대적으로 열세인 예측오차 통계량을 제공하고 있다.

## 4. 결론

본 연구에서는 다중 임계점을 가진 비대칭 금융 시계열 분석을 제안하였으며 특히 가장 간단한 두 개의 임계점을 가진 A-ARCH(1) 점화식을 미국 S&P500지수를 사용하여 자세하게 예시하고 설명하였다. 모수적 붓스트랩을 통해 경쟁 모형의 예측력을 비교 평가하였으며 이를 통해 임계점이  $\delta = 0.3$ 인 CG 방법을 사용한 A-ARCH(1) 모형,  $\delta = 1.0$ 인 CG를 사용한 A-ARCH(1) 모형이 추천되었다. A-ARCH(1) 모형은 임계점  $\delta$ 와  $-\delta$ 에서 변동성 함수가 불연속이므로 모수 추정에 어려움이 있으며, 어려움을 덜기 위해 임계점 델타를 모수로 간주하지 않고 고정된 상수로 취급하여 다양하게 변화시키면서 분석을 수행하였다. 임계점 델타를 모수로 간주하여 다른 모수들과 동시에 추정하는 것은 많은 이론적 어려움이 있을 것으로 판단하며 이는 후후 과제

Table 3: Summary statistics for 1-step ahead prediction error

Model	MSE	MAD	MAPE	WMAPE
<b>B = 300</b>				
TARCH(1)	4.075988	1.192436	111.2315	99.842
A-ARCH(1) NM ( $\delta = 0.3$ )	4.097409	1.197955	111.2302	100.304
A-ARCH(1) NM ( $\delta = 0.5$ )	4.050892	1.191477	113.5525	99.7617
A-ARCH(1) NM ( $\delta = 1.0$ )	4.030512	1.193908	<b>105.8836</b>	99.9652
A-ARCH(1) BFGS ( $\delta = 1.0$ )	4.044270	1.190379	107.6714	99.6698
A-ARCH(1) CG ( $\delta = 0.3$ )	<b>4.025421</b>	<b>1.185438</b>	108.6011	<b>99.256</b>
A-ARCH(1) CG ( $\delta = 1.0$ )	<b>4.020114</b>	<b>1.188005</b>	118.5278	99.4709
A-ARCH(1) SANN ( $\delta = 0.3$ )	4.046292	1.188560	105.2400	99.5174
A-ARCH(1) SANN ( $\delta = 1.0$ )	4.044583	1.192466	117.4325	99.8444
<b>B = 1000</b>				
TARCH(1)	4.067711	1.192695	113.0487	99.8636
A-ARCH(1) NM ( $\delta = 0.3$ )	4.036550	1.188352	101.5959	99.5
A-ARCH(1) NM ( $\delta = 0.5$ )	4.082832	1.197914	103.8931	100.3006
A-ARCH(1) NM ( $\delta = 1.0$ )	4.066807	1.190198	102.2168	99.6545
A-ARCH(1) BFGS ( $\delta = 1.0$ )	4.035890	1.190167	<b>101.2224</b>	99.652
A-ARCH(1) CG ( $\delta = 0.3$ )	<b>4.007491</b>	<b>1.187945</b>	102.2189	99.4659
A-ARCH(1) CG ( $\delta = 1.0$ )	4.072110	1.191361	106.0702	99.752
A-ARCH(1) SANN ( $\delta = 0.3$ )	4.010080	1.189401	106.9835	99.5878
A-ARCH(1) SANN ( $\delta = 1.0$ )	<b>4.009554</b>	<b>1.185418</b>	102.2919	<b>99.2544</b>
<b>B = 3000</b>				
TARCH(1)	4.038612	1.191641	104.3857	99.7754
A-ARCH(1) NM ( $\delta = 0.3$ )	4.067551	1.192463	101.2077	99.8442
A-ARCH(1) NM ( $\delta = 0.5$ )	4.028637	1.192467	<b>100.773</b>	99.8445
A-ARCH(1) NM ( $\delta = 1.0$ )	4.034495	1.190912	102.2726	99.7144
A-ARCH(1) BFGS ( $\delta = 1.0$ )	4.025215	1.190584	105.9692	99.6869
A-ARCH(1) CG ( $\delta = 0.3$ )	4.045526	1.193881	102.2452	99.9629
A-ARCH(1) CG ( $\delta = 1.0$ )	<b>4.016614</b>	<b>1.189276</b>	103.0242	<b>99.5774</b>
A-ARCH(1) SANN ( $\delta = 0.3$ )	4.046580	1.192473	102.4417	99.845
A-ARCH(1) SANN ( $\delta = 1.0$ )	4.044811	1.194614	103.6877	100.0244

남겨두고자 한다. 또한, 3개 이상의 임계점을 가진 고차의 비대칭-ARCH 모형이나 비대칭-GARCH 모형으로 확장하는 연구도 흥미로운 미래 연구주제이다.

## References

- Belisle CJP (1992). Convergence theorems for a class of simulated annealing algorithms, *Journal of Applied Probability*, **29**, 885–895.
- Bollerslev T (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31(3)**, 307–327.
- Byrd RH, Lu P, Nocedal J, and Zhu C (1995). A limited memory algorithm for bound constrained optimization, *SIAM Journal on Scientific Computing*, **16**, 1190–1208.



- Choi SW, Hwang SY, and Lee SD (2020). Volatility-non-stationary GARCH(1,1) models featuring threshold-asymmetry and power transformation, *Korean Journal of Applied Statistics*, **33**, 713-722.
- Choi SW, Yoon JE, Lee SD, and Hwang SY (2021). Asymmetric and non-stationary GARCH(1,1) models: parametric bootstrap to evaluate forecasting performance, *Korean Journal of Applied Statistics*, **34**, 611-621.
- Engle RF (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**, 987–1008.
- Fletcher R and Reeves CM (1964). Function minimization by conjugate gradients. *Computer Journal*, **7**, 148–154.
- Guerrout E, Ait-Aoudia S, Michelucci D, and Mahiou R (2018). Hidden Markov random field model and Broyden–Fletcher–Goldfarb–Shanno algorithm for brain image segmentation, *Journal of Experimental & Theoretical Artificial Intelligence*, **30(3)**, 415-427.
- Hwang SY, Baek JS, Park JA, and Choi MS (2010). Explosive volatilities for threshold-GARCH processes generated by asymmetric innovations, *Statistics & Probability Letters*, **80**, 26-33.
- Kim DR and Hwang SY (2020). Forecasting evaluation via parametric bootstrap for threshold-INARCH models, *Communications for Statistical Applications and Methods*, **27**, 177-187.
- Miguel JA and Olave P (1999). Bootstrapping forecast intervals in ARCH models, *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research*, **8**, 345-364.
- Nelder JA and Mead R (1965). A simplex algorithm for function minimization. *Computer Journal*, **8**, 308–313.
- Nocedal J and Wright SJ (1999). *Numerical Optimization*. Springer.
- Rabemananjara R and Zakoian JM (1993). Threshold ARCH models and asymmetries in volatility, *Journal of Applied Econometrics*, **8**, 31-49.
- Tsay RS (2010). *Analysis of Financial Time Series*, Third Ed, Wiley, New York.
- Yoon JE and Hwang SY (2015). Volatility computations for financial time series : high frequency and hybrid method, *Korean Journal of Applied Statistics*, **28**, 1163-1170.
- Yoon JE, Lee JW, and Hwang SY (2014). News impact curves of volatility for asymmetric GARCH via LASSO, *Korean Journal of Applied Statistics*, **27**, 159-168.

# 비대칭 금융 시계열을 위한 다중 임계점 변동성 모형

이효령<sup>a</sup>, 황선영<sup>1,a</sup>

<sup>a</sup>숙명여자대학교 통계학과

---

## 요약

본 논문에서는 금융 시계열 비대칭 변동성을 모형화하기 위해서 다중 임계점을 가진 비대칭-ARCH 점화식(A-ARCH(1))을 제안하고 있다. 특히 임계점이 두 개인 간단한 모형에 초점을 맞추어 설명하고 있으며 미국 S&P500 자료 분석을 통해 제시하였다. 다양한 A-ARCH(1) 모형의 예측력 비교를 위해 모수적-붓스트랩을 활용하여 예측오차의 평가 및 예측구간의 정확도를 설명하였다.

주요용어: 비대칭 변동성, 다중 임계점, 모수적 붓스트랩

---

<sup>1</sup>교신저자: (04310) 서울특별시 용산구 청파로47길 100, 숙명여자대학교 통계학과. E-mail: shwang@sookmyung.ac.kr