

AR(1) 모형의 모수에 대한 L-추정법 *

한상문¹⁾ 정병철²⁾

요약

본 연구에서는 AR(1) 과정을 따르는 시계열 모형에서 가산적 이상치(Additive Outlier)가 존재하는 경우, 1차 자기상관계수에 대한 로버스트 추정방법으로 Ruppert 와 Carroll (1980)에 의해 회귀모형에서 제안된 L-추정법 형태의 절사최소제곱추정 (PE 추정)방법을 제안하였다. 더불어 X축의 이상치에 대한 비중강하(down-weight)의 방법으로 Mallows의 가중함수를 고려한 유계영향 절사최소제곱 (bounded influence PE, BIPE)추정량을 제안하였으며 모의실험을 통하여 각 추정량의 효율성을 비교하였다. 모의실험 결과, 다양한 자료의 오염률상에서 일반화 LAD추정치를 예비 추정치로 고려한 BIPE(LAD)-추정량의 효율이 좋은것으로 나타났다.

주요용어: AR(1), 모수, L-추정법

1. 서론

이상치가 존재하는 시계열 모형에 대하여 Fox (1972), Denby와 Martin (1979), de Jongh 과 de Wet (1985), Haddad (2000)와 Guo (2000) 등 많은 연구자들의 연구가 있다. Fox (1973)는 1차 자기상관을 갖는 시계열 모형 (AR(1))에서 이상치의 종류를 IO 이상치 (innovation outlier)와 AO 이상치 (additive outlier, AO)로 구분하여 연구하였다. Denby 와 Martin (1979)은 IO와 AO를 갖는 AR(1) 모형에서 1차 자기상관계수에 대한 로버스트(robust) 추정방법에 대하여 연구하였다. 그들은 IO 모형에서는 M-추정량의 효율이 좋으나 AO 모형에서는 M 추정량의 효율이 떨어져 그에 대한 대안으로 일반화 M-(Generalized M, GM) 추정량의 사용을 제안하였다. de Jongh과 de Wet (1985)은 Koenker와 Bassett (1978)이 제안한 회귀 분위수(Regression Quantile, RQ) 추정방법을 이상치가 존재하는 AR(1) 모형에 적용하였으며, 더불어 X축의 이상치에 대한 비중강하(down-weight) 방법으로 Mallows (1973)가중치 함수를 이용한 유계영향(bounded-influence) RQ (BIRQ) 추정방법을 제안하였다.

본 논문에서는 가산적 이상치 (AO)가 존재하는 AR(1) 모형에서 1차 자기상관계수에 대한 로버스트 추정방법으로 Ruppert 와 Carroll (1980)에 의해서 제안된 “예비 추정량(Preliminary Estimator, PE)”을 이용한 절사최소제곱추정(Trimmed Least Squares Estimation) 방법 (이하 PE 추정방법)을 제안하고자 한다. 이는 적절한 예비 추정량 (OLS,

* 본 연구는 2004년 서울시립대학교 학술연구비 지원에 의하여 수행되었음.

1) (136-701) 서울시 동대문구 전농동 90번지 서울시립대학교 통계학과 교수.

E-mail: smhan@uos.ac.kr

2) (136-742) 서울시 성북구 동선동 249-1 성신여자대학교 통계학과 초빙교수.

E-mail: bcjung@empal.com

LAD 등)에 의하여 모형을 적합한 후, 얻어진 잔차의 절대값이 일정 수준보다 큰 관측값을 제거한 후 모수에 대하여 재추정하는 L-추정 형태의 추정량이다. 하지만 PE-추정량은 X 축 방향의 이상치에 대하여 유계되어 있지 못하다는 단점이 존재한다. 이와 같은 문제를 해결하기 위하여 본 논문에서는 X 축의 이상치에 대한 비중강하(down-weight)의 방법으로 de Jongh과 de Wet (1985)이 고려한 Mallows의 가중치 함수를 이용하는 유계영향 PE (이하 BIPE) 추정량을 제안하고자 한다. 더불어 모의실험을 통하여 제안된 L-추정량의 효율성을 비교하고자 한다.

2. 모형 및 PE 추정량

다음과 같은 가산적 이상치 (additive outlier, AO)를 갖는 1차 자기상관을 갖는 모형을 고려해보자.

$$X_t = \mu + \rho X_{t-1} + \varepsilon_t \quad (2.1)$$

$$Y_t = X_t + \nu_t \quad (2.2)$$

여기서 μ 와 ρ 는 각각 상수항과 1차 자기상관계수를 나타내는 모수이며, ε_t 는 $N(0, \sigma_\varepsilon^2)$ 를 따르는 IID 오차항을 나타낸다. 더불어 가산적 이상치를 나타내는 ν_t 는 다음과 같이 “0에서 퇴행”과 “평균이 0인 확률분포”와의 오염된 분포를 따른다고 가정한다.

$$CND(\cdot | \gamma, \sigma^2) = (1 - \gamma)\delta(\cdot) + \gamma F \quad (2.3)$$

γ 는 오염률을 나타내고 $\delta(\cdot)$ 는 0에 퇴행(degenerate)하는 값을 나타내므로 $Pr(\nu_t = 0) = 1 - \gamma$ 이며 $Pr(\nu_t \neq 0) = \gamma$ 이다. 이때 F 는 ν_t 의 오염된 부분에 해당하는 확률분포로 평균이 0이고 대칭인 분포를 나타내며, 정규분포, SLASH, SLACU 및 CAUCHY 등의 분포가 사용될 수 있다. 식 (2.2)와 같은 시계열 모형은 정규분포를 따르는 관측치 X_t 에 “가산적 효과” ν_t 가 더해진 형태로 구성된 시계열 모형이다.

먼저 식 (2.2)의 모형에서 모수 μ 와 ρ 에 대한 로버스트 추정방법으로, Ruppert와 Carroll (1980)에 의하여 제안된 예비 추정치에 의한 절사최소제곱추정(Trimmed Least Squares) 방법 (PE 추정방법)을 적용해보자. $\beta = (\mu \ \rho)'$ 이라 했을 때, Ruppert와 Carroll (1980)의 PE-추정량은 다음과 같은 과정을 통하여 얻을 수 있다.

Step 1. 사전 추정된 회귀계수 $\hat{\beta}_0$ 을 이용하여 다음과 같은 잔차를 얻을 수 있다.

$$r_t = y_t - X_t' \hat{\beta}_0, \quad t = 2, \dots, n, \quad (2.4)$$

여기서 $X_t = (1 \ y_{t-1})'$ 이다.

Step 2. 식 (2.4)에서 얻어진 잔차를 크기순으로 정렬한 후, r_{1n} 과 r_{2n} 을 각각 $[n\alpha]$ 와 $[n(1 - \alpha)]$ 번째 순서에 해당하는 잔차값으로 정의하자. 여기서 α 는 절사율을 나타내며 $[a]$ 는 a 보다 작거나 같은 가장 큰 정수를 나타낸다.

Step 3. 절사를 α 를 갖는 PE-추정량 $\hat{\beta}_{PE}(\alpha)$ 는 다음과 같은 식의 조건을 만족하는 모든 관측치를 제거한 후, 나머지 관측값들만의 최소제곱추정량으로 얻어진다.

$$r_t < r_{1n} \quad \text{or} \quad r_t > r_{2n}. \quad (2.5)$$

Step 1에서 Step 3까지의 과정을 통하여 얻어지는 $\hat{\beta}_{PE}(\alpha)$ 는 행렬과 벡터의 형태로 다음과 같이 재표현 할 수 있다.

$$\hat{\beta}_{PE}(\alpha) = (Z_1' D Z_1)^{-1} Z_1' D Y_2, \quad (2.6)$$

여기서 $Y_2 = (y_2, \dots, y_n)'$ 인 $(n-1) \times 1$ 벡터, $Z_1 = (i_{n-1} \ Y_1)$ 인 $(n-1) \times 2$ 행렬, $Y_1 = (y_1, \dots, y_{n-1})'$ 인 $(n-1) \times 1$ 벡터, i_{n-1} 은 모든 원소가 1로 이루어진 $(n-1) \times 1$ 벡터를 각각 나타내며 D 는 t 번째 원소가 식 (2.5)의 조건을 만족하면 t 번째 대각원소의 값이 0의 값을 갖고 만족하지 못하면 1을 갖는 $(n-1) \times (n-1)$ 대각행렬이다.

식 (2.6)에서 얻어진 $\hat{\beta}_{PE}(\alpha)$ 는 잔차를 얻기 위한 예비 추정치(Preliminary estimate)로 어떤 추정치를 고려했는가에 따라 그 효율성이 달라질 수 있다. 이때 고려할 수 있는 예비 추정치로는 OLS, LAD 또는 RQ-추정치 $\hat{\beta}_0(RQ, \alpha) = [\hat{\beta}_{RQ}(\alpha) + \hat{\beta}_{RQ}(1-\alpha)]/2$ 등을 생각할 수 있다. 이 경우 $\hat{\beta}_{RQ}(\theta)$ 는 θ 번째 RQ 추정치로 다음과 같은 식을 최소화하는 추정치이다.

$$\min_{\beta} \left[\theta \sum_1 |Y_t - X_t' \beta| + (1-\theta) \sum_2 |Y_t - X_t' \beta| \right], \quad (2.7)$$

여기서 $\sum_1 |Y_t - X_t' \beta|$ 은 $Y_t - X_t' \beta > 0$ 인 모든 관측값을 더하는 기호이고, $\sum_2 |Y_t - X_t' \beta|$ 은 $Y_t - X_t' \beta < 0$ 인 모든 관측값을 더하는 기호를 각각 나타낸다. Ruppert와 Carroll (1980)은 선형회귀모형에서 식 (2.7)과 같은 RQ 추정치를 예비 추정치로 고려한 경우 얻어지는 PE-추정량을 $\hat{\beta}_{PE}(RQ, \alpha)$ 라 했을 때, 이는 Koenker와 Bassett (1978)이 제안한 회귀계수에 대한 RQ-추정량과 같은 극한분포를 갖는다는 사실을 보여주었다. 본 연구에서는 PE-추정량의 예비 추정치로 LAD-추정치와 $\hat{\beta}_0(RQ, \alpha)$ 를 고려하였으며, 이때 얻어진 추정량을 각각 PE(LAD)와 PE(RQ)로 정의하자.

3. BI-PE 추정량 제안

그러나 식 (2.6)에서 구해진 $\hat{\beta}_{PE}(\alpha)$ 는 잔차가 큰 부분에 대한 조절을 하기 때문에 Y 축 방향의 이상점에 대하여는 유계되어 있지만 X 축 방향의 이상점에 대해서는 유계되어 있지 않다는 문제점이 존재한다. 이러한 문제를 해결하기 위하여 본 절에서는 X 축 방향의 이상점에 대하여 Mallows의 가중함수를 이용한 유계영향 (bounded influence) PE-추정량 (BIPE-추정량)을 제안하고자 한다.

먼저 예비 추정치로 RQ 추정방법에 X 축 방향의 이상점에 대하여 적절한 가중치를 이용하여 얻는 일반화 RQ-추정치를 이용하는 경우를 고려해보자. 이 경우 $w_t, t = 1, \dots, n-1$ 을 t 번째 개체에 대한 가중치라 했을 때, θ^{th} 일반화 RQ-추정량 $\hat{\beta}_{RQ}^w(\theta)$ 는 다음과 같은 식을 최소화하는 추정량으로 정의된다.

$$\hat{\beta}_{RQ}^w(\theta) = \min_{\beta} \left[\theta \sum_1 w_t |Y_t - X_t' \beta| + (1-\theta) \sum_2 w_t |Y_t - X_t' \beta| \right]. \quad (3.1)$$

그러므로 절사율 α 를 갖는 예비 추정치는 $\widehat{\beta}_0^w(RQ, \alpha) = [\widehat{\beta}_{RQ}^w(\alpha) + \widehat{\beta}_{RQ}^w(1 - \alpha)]/2$ 를 사용할 수 있으며, 이를 이용한 일반화 PE-추정량은 다음과 같이 표현된다.

$$\widehat{\beta}_{PE}^w(RQ, \alpha) = (Z_1' D_\alpha^{(w1)} Z_1)^{-1} Z_1' D_\alpha^{(w1)} Y_2, \quad (3.2)$$

여기서 $D_\alpha^{(w1)}$ 는 $(n-1) \times (n-1)$ 대각행렬로 그의 t 번째 원소인 $d_t^{(w1)}(\alpha)$ 는 다음과 같은 값을 갖게 된다.

$$d_t^{(w1)}(\alpha) = \begin{cases} w_t & \text{if } r_{1n}^w(RQ, \alpha) \leq r_t^w(RQ, \alpha) \leq r_{2n}^w(RQ, \alpha) \\ 0 & \text{otherwise} \end{cases} \quad (3.3)$$

여기서 $r_t^w(RQ, \alpha) = y_t - X_t' \widehat{\beta}_0^w(RQ, \alpha)$ 이며 $r_{1n}^w(RQ, \alpha)$ 와 $r_{2n}^w(RQ, \alpha)$ 는 각각 $[n\alpha]$ 와 $[n(1-\alpha)]$ 번째 순서에 해당하는 잔차이다. 식 (3.2)에서 얻어진 $\widehat{\beta}_{PE}^w(RQ, \alpha)$ 를 표기상 BIPE(RQ)라 놓도록 한다.

다음으로 예비 추정치로 일반화 LAD를 사용하는 일반화 PE-추정량을 고려해보자. 이 경우 예비 추정치로 고려되는 일반화 LAD는 다음과 같이 얻어지는 추정치이다.

$$\widehat{\beta}_0^w(LAD) = \min_{\beta} \left[\theta \sum_{t=2}^n w_t |Y_t - X_t' \beta| \right]. \quad (3.4)$$

식 (3.4)와 같이 얻어지는 예비 추정치를 사용한 일반화 PE-추정량은 다음과 같은 형태를 갖게 된다.

$$\widehat{\beta}_{PE}^w(LAD, \alpha) = (Z_1' D_\alpha^{(w2)} Z_1)^{-1} Z_1' D_\alpha^{(w2)} Y_2, \quad (3.5)$$

여기서 $D_\alpha^{(w2)}$ 는 $(n-1) \times (n-1)$ 대각행렬로 그의 t 번째 원소인 $d_t^{(w2)}(\alpha)$ 는 다음과 같은 값을 갖게 된다.

$$d_t^{(w2)}(\alpha) = \begin{cases} w_t & \text{if } r_{1n}^w(LAD) \leq r_t^w(LAD) \leq r_{2n}^w(LAD) \\ 0 & \text{otherwise} \end{cases} \quad (3.6)$$

여기서 $r_t^w(LAD) = y_t - X_t' \widehat{\beta}_0^w(LAD)$ 이며 $r_{1n}^w(LAD, \alpha)$ 와 $r_{2n}^w(LAD, \alpha)$ 는 각각 $[n\alpha]$ 와 $[n(1-\alpha)]$ 번째 순서에 해당하는 잔차이다. 식 (3.5)에서 얻어진 $\widehat{\beta}_{PE}^w(LAD, \alpha)$ 를 표기상 BIPE(LAD)라 놓도록 한다.

4. 모의실험

본 논문에서 제안한 PE-추정량(PE(LAD), PE(RQ))과 유계영향 PE-추정량(BIPE(LAD), BIPE(RQ))의 효율성을 알아보기 위하여 모의실험을 실시하였다. 모의실험은 Denby와 Martin (1979) 및 de Jongh과 de Wet (1985)에서 사용한 방법과 유사한 방법을 사용하였다. 모의실험은 AO 모형만을 대상으로 하였으며, 표본수 n 은 100으로 고정하였다. 상수

항 μ 는 0으로 고정하였고 1차 자기상관계수 ρ 는 0.5과 0.8을 사용하였다. 또한 식 (2.3)의 가산적 오차항 부분에 나타나는 오염률 γ 의 값은 0에서 0.2까지 0.05단위씩 증가시켰으며, 가산적 오차항 ν_t 의 확률분포 F 로는 정규분포를 사용하였다. 이때 ν_t 의 분산을 9, 36, 100 및 ∞ (SLASH 분포)으로 변화시켜가며 실험하였다.

모의실험에서는 OLS추정량, de Jongh과 de Wet (1985)이 고려한 유계영향 RQ(BIRQ) 추정량 및 Huber-형태의 GM 추정량을 본 연구에서 제안한 추정량과 효율성을 비교하기 위하여 사용하였다. BIPE-추정량 및 BIRQ-추정량을 구하기 위하여 필요한 절사율 α 의 값은 0.05, 0.1 및 0.15등을 사용하였으며 X 축방향의 이상치에 대한 비증강하의 목적으로 Mallows의 가중함수가 사용되었다. 즉, $Y_{(1)} \leq Y_{(2)} \leq \dots \leq Y_{(n-1)}$ 을 $Y_t, t = 1, \dots, n-1$ 들의 순서화된 값이라 하고 z_1, \dots, z_{n-1} 을 Y_t 의 순위라 했을 때 $L = \lceil \tau n + 1 \rceil$ 과 $U = n - L$ 을 각각 정의하자. 본 연구에서는 $\tau = 0.1$ 로 고정하였다. 이와 같은 정의하에서, $(n-1)$ 개의 관측치 Y_1, \dots, Y_{n-1} 에 대하여 다음과 같은 Mallows형태의 가중치를 고려하였다.

$$w_t = \begin{cases} 1 & \text{if } L \leq z_t \leq U \\ \frac{Y_{(L)} - Y_{(U)}}{2Y_{(L-1)} - Y_{(U)} - Y_{(L)}} & \text{if } z_t < L \\ \frac{Y_{(U)} - Y_{(L)}}{2Y_{(U-1)} - Y_{(U)} - Y_{(L)}} & \text{if } z_t > U \end{cases} \quad (4.1)$$

모든 ρ, γ 및 ν_t 의 분산수준에 대하여 1,000번의 반복이 실시되었으며, 각 반복마다 μ 와 ρ 에 대한 추정량을 계산하였다. 이와 같이 얻어진 1,000개의 추정량을 이용하여 각 추정량의 평균제곱오차(Mean Square Error, MSE) 및 편의(Bias)를 계산하였다.

4.2. MSE 비교

표 4.1과 4.2는 $\rho = 0.5, 0.8$ 인 경우 고려된 가산적 이상치의 분산과 오염률 (γ)에 대하여 OLS, 절사율 $\alpha = 0.05, 0.1$ 및 0.15을 고려한 BIPE-(BIPE(LAD), BIPE(RQ)) 추정량과 BIRQ-추정량 및 Huber-형태의 GM 추정량들에 MSE를 나타낸다.

먼저 OLS 추정량은 자료의 오염률이 $\gamma = 0$ 인 경우 기대했던 대로 고려된 다른 추정량들보다 작은 MSE를 보여주고 있다. 반면 오염률 $\gamma \geq 0.05$ 인 경우 다른 모든 추정량들에 비하여 큰 MSE를 보이고 있다. X 축 방향의 이상치에 대한 비증강하를 고려하지 않은 PE-추정량(PE(LAD), PE(RQ))은 오염률 $\gamma \geq 0.05$ 인 모든 경우 X 축 방향의 이상치에 대한 비증강하를 고려한 추정량인 BIRQ-추정량, BIPE-추정량 및 Huber-형태의 GM추정량에 비하여 큰 MSE 수준을 보여주고 있어 AO 모형에서는 효율적이지 못하였다. Huber-형태의 GM 추정량은 참 자기상관계수가 $\rho = 0.5$ 인 경우 L-추정량과 거의 비슷한 효율을 보이고 있는 반면 $\rho = 0.8$ 인 경우 BIPE-추정량에 비하여 큰 MSE를 보여주고 있다. 특히 $\rho = 0.8$ 인 경우 자료의 오염률이 0보다 큰 경우 오염률 수준에 관계없이 가장 좋은 효율을 보이는 BIPE-추정량에 비하여 대략 50% 이상 큰 MSE를 보이고 있다. de Jongh과 de Wet (1985)이 제안한 BIRQ-추정량과 본 연구에서 제안한 예비 추정치를 일반화 RQ추정량으로 사용한

BIPE(RQ) 추정량도 Huber-형태의 GM 추정량과 비슷한 결과를 보여주고 있다. 다만 이들은 $\rho = 0.8$ 이고 오염률이 0보다 큰 경우 가장 좋은 효율을 보이는 추정량에 비하여 최소 50%에서 최대 3배이상 큰 MSE를 보이는 보여 Huber형태의 GM-추정량보다 효율이 떨어지고 있다. 반면, 예비 추정치를 일반화 LAD를 사용한 BIPE(LAD)-추정량의 효율은 오염률이 0보다 큰 경우 전반적으로 가장 좋은 것으로 나타났다. 특히 $\gamma = 0.05$ 인 경우 ν_t 의 분산 크기에 관계없이 절사율 $\alpha = 0.05$ 를 고려한 BIPE(LAD) 추정량의 효율이 좋은 것으로 나타났으며, 오염률이 커질수록 절사율 $\alpha = 0.10$ 또는 0.15 를 사용하는 BIPE(LAD)-추정량의 효율이 좋은 것으로 나타났다.

표 4.1과 4.2의 결과를 이용하면, 가산적 이상치가 존재하는 AR(1) 모형에서 자기상관 계수에 대한 추정량은 오염률이 크지 않는 경우 BIPE(LAD)-추정량이 다른 추정량들에 비하여 약간 효율적이었으며, 오염률이 큰 경우에는 절사율 0.10 또는 0.15인 BIPE(LAD)-추정량의 MSE가 GM 추정량 및 다른 추정량들에 비하여 작게 나타나 이의 사용이 좋을 것으로 생각된다. 더불어 이와 같은 효율성의 차이는 자기상관이 강할수록 더 심해지는 것으로 나타났다.

4.2. 편의 비교

표 4.3과 4.4는 $\rho = 0.5, 0.8$ 인 경우 고려된 ν_t 의 분산과 오염률 (γ)에 대하여 ρ 에 대한 다양한 추정량들의 편의 (bias)를 나타낸다. 먼저 고려된 모든 추정량들의 편의값은 음(-)의 값으로 나타나 각 추정량들은 참 모수 ρ 값을 과소추정하고 있음을 보여준다. 특히 OLS추정량은 오염률이 0인 경우를 제외하고 다른 추정량들의 편이에 비하여 아주 큰 편의를 보이고 있다. 더불어 각 추정량의 편이는 오염률 γ 값이 커질수록 증가하는 경향을 보이고 있다.

표 4.3과 4.4에 따르면 ρ 의 값과 오염률 γ 값에 관계없이 예비 추정치를 일반화 LAD로 사용한 BIPE(LAD)-추정량의 편이가 작게 나타나고 있음을 볼 수 있다. 특히 절사율 $\alpha = 0.1$ 을 사용한 BIPE(LAD)-추정량은 ν_t 의 분산이 100보다 작거나 같고 오염률 γ 가 0.15보다 작은 경우 다른 추정량들에 비해 작은 편의를 보이고 있으며, 오염률이 0.15 및 0.20인 경우에는 절사율 $\alpha = 0.1$ 을 사용한 BIPE(LAD)-추정량이 가장 작은 편의를 보이고 있다. 반면 ν_t 의 분산이 ∞ 인 SLASH처럼 무거운 꼬리를 갖는 분포를 따르는 경우, 오염률 $\gamma \leq 0.05$ 인 경우 절사율 $\alpha = 0.05$ 를 사용한 BIPE(LAD)-추정량의 편이가 가장 작게 나타났으며 $\gamma \geq 0.10$ 인 경우 절사율 $\alpha = 0.1$ 을 사용한 BIPE(LAD)-추정량의 편이가 가장 작게 나타났다.

4.3. 음(-)의 자기상관에 대한 추정

음(-)의 자기상관이 존재하는 AO 모형에서 각 추정량들의 효율을 알아보하고자 1차 자기상관계수가 -0.8과 -0.5인 경우에 대한 모의실험을 실시하였다. 이때 ρ 값을 제외한 모든 모의실험 방법은 앞 절과 동일하게 사용하였다. 다음 표 4.5 과 표 4.6은 가산적 오차항 ν_t 의 분산이 36인 경우 각 추정량의 MSE와 편의를 나타낸다. ν_t 의 분산이 9 및 100인 경우와 ν_t 가 slash분포를 따르는 경우에도 비슷한 결과를 보여주어 지면관계상 신지않았다.

표 4.5에서 v_t 가 $N(0, 36)$ 일 경우 음(-)의 자기상관계수에 대한 각 추정량들의 MSE는 양(+)의 자기상관계수에 대한 추정량들과 비슷한 결과를 보여주고 있다. 즉, 음(-)의 자기상관에 대해서도 오염률 γ 값에 따라 절사율이 10% 또는 15%를 고려한 BIPE(LAD)추정량의 효율성이 가장 좋은것으로 나타났다.

반면 추정량의 편의(bias)를 나타내는 표 4.6의 결과를 살펴보면 음(-)의 자기상관계수에 대해서는 양(+)의 자기상관계수인 경우의 결과인 표 4.3과 표 4.4와 달리 모든 추정량들이 양(+)의 편의를 나타내어 참 모수 ρ 를 과대추정하고 있음을 보여준다. 각 추정량의 bias에 대한 비교에 있어서도 오염률 γ 값에 따라 절사율이 10% 또는 15%를 고려한 BIPE(LAD)추정량의 편위가 가장 작게 나타났다.

5. 결론

본 연구에서는 가산적 이상치를 갖는 AR(1) 모형에서 예비 추정량을 이용한 절사 최소제곱추정 (PE 추정) 방법에 대하여 연구하였다. 이는 다중회귀모형에 대한 Ruppert 와 Carroll (1980)의 연구를 시계열 모형으로 확장한것이다. 하지만 AR(1) 모형에서 PE 추정량은 X축 방향의 이상치에 유계되어 있지못하다는 단점이 존재하므로 X축의 이상치에 대한 비중강하(down-weight)의 방법으로 Mallows의 가중함수를 고려한 유계영향 절사 최소제곱추정량 (BIPE-추정량)을 제안하였다. 모의실험 결과, 오염률의 수준에 따라 절사율 α 를 0.05, 0.1 및 0.15로 조절한 일반화 LAD를 예비 추정치로 사용하는 BIPE(LAD)-추정량의 사용이 가장 효율적이었다.

참고문헌

- De Jongh, P.J. and De Wet T. (1985). Trimmed mean and bounded influence estimators for the parameters of the AR(1) process, *Communications in Statistics -Theory and Methods*, **14**, 1361-1375.
- Denby, L. and Martin, R.D. (1979). Robust estimation of the first-order autoregressive parameter, *Journal of the American Statistical Association*, **74**, 140-146.
- Fox, A.J. (1972). Outliers in time series. *Journal of the Royal Statistical Society, Series B*, **34**, 350-363.
- Guo, J-H. (2000). Robust estimation for the coefficient if a first order autoregressive process, *Communications in Statistics -Theory and Methods*, **29**, 55-66.
- Haddad, J.N. (2000). On robust estimation in the first order autoregressive processes, *Communications in Statistics -Theory and Methods*, **29**, 45-54.
- Koenker, R. and Bassett, G. (1978). Regression quantiles, *Econometrica*, **46**, 33-50.
- Mallows, C.L. (1973). Influence functions, *Unpublished paper presented at a conference on robust regression held at Cambridge, Mass., and sponsored by the National Bureau of Economic Research*.
- Ruppert, D. and Carroll, R. (1980). Trimmed least squares estimation in the linear model, *Journal of the American Statistical Association*, **75**, 828-838.

표 4.1: $\rho = 0.5$ 인 경우 오염률(γ), 절사율(α)에 따른 각 추정량들의 MSE

ρ	F	오염률 (γ)	OLS	절사율($\alpha = 0.05$)						절사율($\alpha = 0.10$)						절사율($\alpha = 0.15$)						Huber (GM)
				PE (RQ)	PE (LAD)	BIRQ (RQ)	BIPE (LAD)	BIPE (RQ)	BIPE (LAD)	PE (RQ)	PE (LAD)	BIRQ (RQ)	BIPE (LAD)	BIPE (RQ)	BIPE (LAD)	PE (RQ)	PE (LAD)	BIRQ (RQ)	BIPE (LAD)	BIPE (RQ)	BIPE (LAD)	
0.5	N(0, 9)	0.00	<u>0.008</u>	0.009	0.012	0.010	0.010	0.013	0.010	0.014	0.010	0.010	0.016	0.010	0.016	0.010	0.010	0.010	0.010	0.018	0.010	
		0.05	0.032	0.027	0.024	0.018	0.017	0.017	0.017	0.027	0.026	0.017	0.020	0.020	0.026	0.028	0.017	0.017	0.017	0.022	<u>0.016</u>	
		0.10	0.057	0.050	0.044	0.031	0.030	0.025	0.025	0.049	0.045	0.029	0.027	0.027	0.049	0.048	0.028	0.027	0.027	0.030	0.026	
		0.15	0.082	0.075	0.068	0.048	0.046	<u>0.039</u>	0.039	0.073	0.066	0.045	0.043	<u>0.039</u>	0.072	0.068	0.043	0.042	0.042	0.041	0.040	
		0.20	0.103	0.096	0.088	0.067	0.056	0.056	0.094	0.086	0.063	0.062	<u>0.053</u>	0.093	0.087	0.061	0.059	0.055	0.055	0.055		
		0.00	<u>0.008</u>	0.009	0.012	0.009	0.009	0.013	0.010	0.015	0.009	0.010	0.016	0.010	0.016	0.010	0.010	0.010	0.010	0.017	0.009	
		0.05	0.090	0.069	0.058	0.029	0.028	<u>0.021</u>	0.021	0.068	0.058	0.026	0.024	0.022	0.067	0.060	0.024	0.023	0.024	0.024	0.023	
		0.10	0.144	0.123	0.111	0.057	0.057	0.039	0.123	0.110	0.052	0.049	0.036	0.120	0.111	0.048	0.045	0.038	0.038	0.044	0.044	
	N(0, 36)	0.15	0.177	0.159	0.153	0.090	0.091	0.076	0.158	0.149	0.086	0.086	<u>0.065</u>	0.158	0.152	0.083	0.080	0.066	0.066	0.073		
		0.20	0.194	0.182	0.176	0.123	0.124	0.113	0.179	0.172	0.118	0.118	0.099	0.178	0.173	0.116	0.112	<u>0.096</u>	0.099	0.099		
		0.00	<u>0.009</u>	0.010	0.013	0.010	0.010	0.014	0.010	0.015	0.010	0.010	0.016	0.010	0.016	0.010	0.010	0.011	0.011	0.017	0.010	
		0.05	0.150	0.111	0.103	0.037	0.037	<u>0.022</u>	0.112	0.103	0.032	0.029	0.024	0.113	0.107	0.029	0.027	0.026	0.026	0.027	0.027	
	N(0, 100)	0.10	0.198	0.175	0.170	0.080	0.079	0.057	0.175	0.169	0.073	0.068	<u>0.045</u>	0.176	0.172	0.067	0.060	0.048	0.048	0.061		
		0.15	0.223	0.207	0.203	0.124	0.128	0.111	0.202	0.199	0.117	0.117	<u>0.084</u>	0.204	0.201	0.112	0.107	<u>0.084</u>	0.097	0.097		
		0.20	0.238	0.223	0.218	0.165	0.166	0.155	0.218	0.215	0.154	0.155	0.133	0.219	0.215	0.152	0.149	<u>0.128</u>	<u>0.128</u>	0.128		
		0.00	<u>0.008</u>	0.009	0.012	0.010	0.010	0.013	0.010	0.015	0.010	0.010	0.016	0.010	0.016	0.010	0.010	0.011	0.011	0.018	0.010	
	SLASH	0.05	0.074	0.050	0.048	0.018	0.016	<u>0.015</u>	0.049	0.051	0.016	0.018	0.018	0.052	0.054	0.016	0.016	0.019	<u>0.015</u>	0.015		
		0.10	0.128	0.092	0.092	0.034	0.030	<u>0.022</u>	0.092	0.093	0.028	0.026	0.024	0.093	0.096	0.026	0.025	0.025	0.026	0.025		
		0.15	0.162	0.128	0.131	0.053	0.047	<u>0.031</u>	0.131	0.131	0.045	0.041	0.033	0.133	0.135	0.041	0.039	0.036	0.038	0.038		
		0.20	0.190	0.159	0.161	0.075	0.066	0.044	0.159	0.161	0.065	0.058	<u>0.043</u>	0.161	0.162	0.059	0.054	0.046	0.046	0.053		

표 4.2: $\rho = 0.8$ 인 경우 오염률(γ), 절사율(α)에 따른 각 추정량들의 MSE

ρ	F	오염률 (γ)	절사율($\alpha = 0.05$)			절사율($\alpha = 0.10$)			절사율($\alpha = 0.15$)			Huber (GM)							
			PE (RQ)	PE (LAD)	BIQE (RQ)	BIQE (LAD)	PE (RQ)	PE (LAD)	BIRQ (RQ)	BIRQ (LAD)	BIPE (RQ)		BIPE (LAD)						
0.8		0.00	0.006	0.008	0.006	0.006	0.008	0.006	0.009	0.006	0.010	0.006	0.010	0.006					
		0.05	0.036	0.026	0.016	0.018	0.016	0.012	0.024	0.018	0.015	0.014	0.022	0.020	0.014	0.016			
		0.10	0.070	0.057	0.034	0.035	0.033	0.020	0.050	0.032	0.029	0.027	0.046	0.034	0.027	0.026	0.028		
		0.15	0.105	0.091	0.057	0.059	0.056	0.034	0.081	0.048	0.049	0.046	0.074	0.051	0.044	0.042	0.030	0.045	
		0.20	0.136	0.123	0.094	0.085	0.080	0.058	0.116	0.081	0.076	0.071	0.110	0.081	0.068	0.065	0.047	0.068	
		0.00	0.006	0.006	0.008	0.007	0.006	0.008	0.006	0.009	0.009	0.007	0.009	0.007	0.010	0.007	0.007	0.010	0.006
N(0, 36)		0.05	0.143	0.093	0.031	0.031	0.029	0.076	0.032	0.023	0.021	0.013	0.067	0.037	0.020	0.019	0.016	0.022	
		0.10	0.245	0.206	0.111	0.087	0.085	0.040	0.180	0.092	0.064	0.058	0.161	0.097	0.053	0.049	0.029	0.055	
		0.15	0.323	0.295	0.220	0.155	0.155	0.094	0.273	0.185	0.128	0.122	0.256	0.188	0.108	0.101	0.053	0.104	
		0.20	0.374	0.355	0.310	0.224	0.220	0.166	0.343	0.265	0.198	0.193	0.328	0.263	0.175	0.166	0.096	0.160	
		0.00	0.006	0.006	0.008	0.006	0.006	0.008	0.007	0.009	0.006	0.006	0.009	0.006	0.010	0.006	0.006	0.010	0.006
		0.05	0.271	0.178	0.065	0.043	0.040	0.014	0.154	0.066	0.027	0.023	0.014	0.134	0.075	0.023	0.021	0.017	0.025
N(0, 100)		0.10	0.405	0.346	0.242	0.132	0.137	0.326	0.218	0.094	0.082	0.027	0.305	0.227	0.072	0.062	0.030	0.071	
		0.15	0.480	0.440	0.388	0.234	0.237	0.154	0.427	0.349	0.190	0.185	0.417	0.355	0.157	0.141	0.062	0.140	
		0.20	0.521	0.498	0.478	0.335	0.338	0.276	0.489	0.442	0.302	0.304	0.484	0.441	0.275	0.265	0.153	0.234	
		0.00	0.006	0.007	0.008	0.007	0.007	0.008	0.007	0.009	0.007	0.007	0.007	0.007	0.010	0.007	0.007	0.010	0.006
		0.05	0.147	0.063	0.055	0.015	0.013	0.010	0.059	0.058	0.013	0.011	0.012	0.062	0.063	0.011	0.012	0.013	0.012
		0.10	0.245	0.141	0.104	0.033	0.028	0.015	0.129	0.108	0.024	0.023	0.016	0.128	0.114	0.022	0.021	0.019	0.023
SLASH		0.15	0.330	0.219	0.165	0.067	0.060	0.207	0.170	0.047	0.043	0.026	0.203	0.179	0.041	0.040	0.029	0.043	
		0.20	0.394	0.285	0.227	0.099	0.089	0.040	0.275	0.226	0.068	0.062	0.276	0.237	0.058	0.054	0.036	0.060	

표 4.3: $\rho = 0.5$ 인 경우 오염률(γ), 절사율(α)에 따른 각 추정량들의 추정된 BIAS

ρ	F	오염률 (γ)	OLS	절사율($\alpha = 0.05$)						절사율($\alpha = 0.10$)						절사율($\alpha = 0.15$)						Huber (GM)
				PE (RQ)	PE (LAD)	BIRQ (RQ)	BIPE (LAD)	BIPE (RQ)	BIPE (LAD)	PE (RQ)	PE (LAD)	BIRQ (RQ)	BIPE (RQ)	BIPE (LAD)	BIPE (RQ)	BIPE (LAD)	PE (RQ)	PE (LAD)	BIRQ (RQ)	BIPE (RQ)	BIPE (LAD)	
0.5	N(0, 9)	0.00	-0.029	-0.030	-0.028	-0.033	-0.028	-0.027	-0.029	-0.029	-0.031	-0.027	-0.027	-0.029	-0.030	-0.029	-0.030	-0.028	-0.029	-0.029	-0.027	-0.027
		0.05	-0.139	-0.119	-0.096	-0.086	-0.080	-0.063	-0.119	-0.097	-0.082	-0.076	-0.063	-0.116	-0.100	-0.079	-0.074	-0.063	-0.074	-0.074	-0.074	-0.074
		0.10	-0.210	-0.191	-0.165	-0.142	-0.132	-0.109	-0.187	-0.162	-0.134	-0.126	-0.105	-0.186	-0.168	-0.129	-0.123	-0.107	-0.122	-0.122	-0.122	-0.122
		0.15	-0.263	-0.249	-0.230	-0.191	-0.182	-0.159	-0.244	-0.220	-0.183	-0.175	-0.148	-0.245	-0.221	-0.177	-0.169	-0.148	-0.167	-0.167	-0.167	-0.167
		0.20	-0.301	-0.289	-0.274	-0.236	-0.230	-0.206	-0.287	-0.267	-0.226	-0.221	-0.193	-0.285	-0.268	-0.222	-0.216	-0.192	-0.209	-0.209	-0.209	-0.209
N(0, 36)	N(0, 36)	0.00	-0.025	-0.023	-0.025	-0.029	-0.022	-0.023	-0.026	-0.022	-0.026	-0.024	-0.021	-0.023	-0.023	-0.023	-0.026	-0.022	-0.023	-0.023	-0.023	-0.023
		0.05	-0.268	-0.221	-0.181	-0.130	-0.116	-0.075	-0.221	-0.182	-0.117	-0.102	-0.072	-0.219	-0.187	-0.108	-0.102	-0.075	-0.106	-0.106	-0.106	-0.106
		0.10	-0.360	-0.328	-0.304	-0.212	-0.205	-0.149	-0.330	-0.298	-0.196	-0.182	-0.129	-0.325	-0.300	-0.186	-0.174	-0.131	-0.176	-0.176	-0.176	-0.176
		0.15	-0.406	-0.385	-0.377	-0.281	-0.280	-0.246	-0.386	-0.370	-0.273	-0.267	-0.216	-0.384	-0.372	-0.265	-0.258	-0.215	-0.248	-0.248	-0.248	-0.248
		0.20	-0.429	-0.418	-0.412	-0.336	-0.337	-0.318	-0.415	-0.406	-0.330	-0.327	-0.288	-0.414	-0.406	-0.326	-0.317	-0.280	-0.299	-0.299	-0.299	-0.299
N(0, 100)	N(0, 100)	0.00	-0.027	-0.028	-0.030	-0.033	-0.027	-0.030	-0.027	-0.028	-0.030	-0.025	-0.029	-0.027	-0.029	-0.031	-0.027	-0.029	-0.029	-0.029	-0.029	-0.029
		0.05	-0.361	-0.294	-0.268	-0.152	-0.142	-0.078	-0.300	-0.270	-0.134	-0.118	-0.076	-0.304	-0.279	-0.123	-0.112	-0.084	-0.120	-0.120	-0.120	-0.120
		0.10	-0.433	-0.406	-0.398	-0.263	-0.256	-0.196	-0.406	-0.395	-0.244	-0.228	-0.153	-0.409	-0.400	-0.230	-0.209	-0.157	-0.217	-0.217	-0.217	-0.217
		0.15	-0.462	-0.449	-0.446	-0.337	-0.341	-0.310	-0.445	-0.441	-0.327	-0.321	-0.248	-0.447	-0.443	-0.316	-0.303	-0.246	-0.294	-0.294	-0.294	-0.294
		0.20	-0.477	-0.466	-0.464	-0.393	-0.394	-0.380	-0.463	-0.461	-0.383	-0.381	-0.340	-0.465	-0.461	-0.378	-0.370	-0.328	-0.345	-0.345	-0.345	-0.345
SLASH	SLASH	0.00	-0.027	-0.025	-0.025	-0.032	-0.025	-0.024	-0.027	-0.027	-0.028	-0.026	-0.026	-0.027	-0.028	-0.028	-0.026	-0.028	-0.028	-0.028	-0.028	-0.028
		0.05	-0.206	-0.153	-0.134	-0.084	-0.071	-0.052	-0.157	-0.140	-0.075	-0.069	-0.055	-0.161	-0.146	-0.072	-0.067	-0.056	-0.068	-0.068	-0.068	-0.068
		0.10	-0.316	-0.253	-0.239	-0.142	-0.125	-0.089	-0.256	-0.243	-0.126	-0.117	-0.089	-0.260	-0.248	-0.119	-0.112	-0.092	-0.115	-0.115	-0.115	-0.115
		0.15	-0.376	-0.324	-0.318	-0.196	-0.178	-0.128	-0.330	-0.319	-0.177	-0.166	-0.126	-0.334	-0.325	-0.167	-0.158	-0.132	-0.160	-0.160	-0.160	-0.160
		0.20	-0.418	-0.376	-0.374	-0.246	-0.224	-0.170	-0.376	-0.372	-0.226	-0.209	-0.159	-0.380	-0.374	-0.212	-0.200	-0.162	-0.201	-0.201	-0.201	-0.201

표 4.5: ν_t 가 $N(0, 36)$ 을 따르는 경우, 음의 자기상관계수에 대한 각 추정량들의 MSE

ρ	오염률 (γ)	OLS	절사율($\alpha = 0.05$)						절사율($\alpha = 0.10$)						절사율($\alpha = 0.15$)						Huber (GM)
			PE (RQ)	PE (LAD)	BIRQ (RQ)	BIPE (LAD)	BIPE (RQ)	BIPE (LAD)	PE (RQ)	PE (LAD)	BIRQ (RQ)	BIPE (LAD)	BIPE (RQ)	BIPE (LAD)	PE (RQ)	PE (LAD)	BIRQ (RQ)	BIPE (LAD)	BIPE (RQ)	BIPE (LAD)	
-0.5	0.00	0.007	0.008	0.011	0.008	0.009	0.012	0.008	0.014	0.008	0.009	0.015	0.009	0.015	0.009	0.015	0.009	0.009	0.016	0.008	
	0.05	0.080	0.062	0.052	0.022	0.023	0.018	0.060	0.052	0.020	0.020	0.019	0.060	0.054	0.019	0.018	0.020	0.018	0.020	0.018	
	0.10	0.131	0.116	0.111	0.050	0.051	0.039	0.115	0.104	0.045	0.044	0.034	0.113	0.106	0.041	0.039	0.035	0.037	0.037	0.037	
	0.15	0.158	0.154	0.152	0.083	0.086	0.074	0.152	0.145	0.078	0.078	0.061	0.151	0.144	0.074	0.072	0.060	0.065	0.065	0.065	
	0.20	0.177	0.171	0.172	0.110	0.113	0.105	0.171	0.165	0.106	0.109	0.091	0.171	0.165	0.104	0.103	0.089	0.089	0.089	0.089	
-0.8	0.00	0.004	0.005	0.006	0.004	0.005	0.006	0.005	0.007	0.004	0.005	0.007	0.005	0.008	0.005	0.005	0.008	0.004	0.004	0.004	
	0.05	0.112	0.073	0.024	0.022	0.022	0.011	0.051	0.022	0.015	0.015	0.010	0.046	0.026	0.013	0.014	0.013	0.014	0.014	0.014	
	0.10	0.209	0.177	0.103	0.067	0.071	0.037	0.148	0.074	0.049	0.045	0.020	0.128	0.076	0.039	0.036	0.021	0.037	0.037	0.037	
	0.15	0.264	0.246	0.187	0.121	0.127	0.079	0.223	0.143	0.097	0.094	0.044	0.205	0.141	0.080	0.075	0.039	0.073	0.073	0.073	
	0.20	0.331	0.320	0.281	0.187	0.194	0.148	0.303	0.238	0.169	0.168	0.100	0.292	0.229	0.150	0.143	0.085	0.130	0.130	0.130	

표 4.6: ψ_1 가 $N(0, 36)$ 을 따르는 경우, 음의 자기상관계수에 대한 각 추정량들의 추정된 BIAS

ρ	오염률 (γ)	절사율($\alpha = 0.05$)			절사율($\alpha = 0.10$)			절사율($\alpha = 0.15$)			Huber (GM)					
		PE (RQ)	PE (LAD)	BIPE (RQ)	BIPE (LAD)	PE (RQ)	PE (LAD)	BIRQ (RQ)	BIRQ (LAD)	BIPE (RQ)		BIPE (LAD)				
-0.5	0.00	0.000	0.001	-0.000	-0.002	-0.001	0.001	0.002	0.002	0.002	0.002	0.002	0.000	-0.003		
	0.05	0.247	0.205	0.162	0.098	0.097	0.058	0.201	0.160	0.166	0.087	0.079	0.077	0.054	0.079	
	0.10	0.342	0.321	0.303	0.194	0.197	0.154	0.317	0.289	0.180	0.173	0.122	0.168	0.161	0.160	
	0.15	0.383	0.382	0.378	0.267	0.274	0.247	0.380	0.365	0.258	0.254	0.205	0.249	0.242	0.198	
	0.20	0.408	0.405	0.409	0.316	0.322	0.309	0.407	0.398	0.312	0.316	0.277	0.307	0.303	0.266	0.281
-0.8	0.00	0.013	0.015	0.013	0.006	0.011	0.011	0.012	0.012	0.007	0.011	0.010	0.013	0.012	0.011	0.011
	0.05	0.289	0.214	0.090	0.102	0.103	0.049	0.174	0.083	0.079	0.079	0.040	0.164	0.097	0.072	0.074
	0.10	0.432	0.386	0.257	0.223	0.229	0.143	0.340	0.198	0.180	0.170	0.087	0.312	0.208	0.157	0.151
	0.15	0.495	0.475	0.389	0.320	0.328	0.238	0.445	0.316	0.278	0.269	0.154	0.421	0.314	0.245	0.235
	0.20	0.561	0.552	0.506	0.410	0.420	0.352	0.534	0.449	0.386	0.382	0.266	0.521	0.437	0.357	0.344

L-Estimation for the Parameter of the AR(1) Model *

Sang Moon Han ¹⁾ Byoung Cheol Jung ²⁾

ABSTRACT

In this study, a robust estimation method for the first-order autocorrelation coefficient in the time series model following AR(1) process with additive outlier(AO) is investigated. We propose the L-type trimmed least squares estimation method using the preliminary estimator (PE) suggested by Ruppert and Carroll (1980) in multiple regression model. In addition, using Mallows' weight function in order to down-weight the outlier of X-axis, the bounded-influence PE (BIPE) estimator is obtained and the mean squared error (MSE) performance of various estimators for autocorrelation coefficient are compared using Monte Carlo experiments. From the results of Monte-Carlo study, the efficiency of BIPE(LAD) estimator using the generalized-LAD to preliminary estimator performs well relative to other estimators.

Keywords: AR(1), Parameter, L-estimation

* This work was supported by research fund of University of Seoul

1) Professor, Dept. of Statistics, University of Seoul, 90 cheonnong-dong, dongdaemun-gu, Seoul 130-743, Korea,

E-mail : smhan@uos.ac.kr

2) Invited Professor, Dept. of Statistics, Sungshin University, 249-1 dongsun-dong, sungbuk-gu, Seoul 136-742, Korea,

E-mail : bcjung@empal.com