



A New Unbiased Estimator of MLR Model Coefficients Based on AC

Sajjad Piradl* 

Department of Statistics, Payame Noor University, Tehran, Iran

Article information

Article history:

Received: October 10, 2025

Revised: December 28, 2025

Accepted: January 14, 2026

Available online: June 1, 2026

Keywords:

MLR model; Unbiased estimator; AC-based estimator; LS estimator; ARMA model.

Correspondence:

Sajjad Piradl*

sajjadpiradl@pnu.ac.ir

Abstract

One of the well-known statistical methods in predictive analysis is the use of the multiple linear regression (MLR) model. In many studies, various estimators have been proposed to estimate the coefficients of this type of regression model, of which the ordinary least squares (OLS) estimator is one of the most famous and, at the same time, is one of the most accurate. This paper introduces a new estimator of MLR model coefficients based on autocovariance (AC). It is shown that although the AC-based estimator proposed in this paper may not be intuitively appealing, it is an unbiased estimator of the model coefficients. It is also shown that if the vector of independent variables satisfies certain regularity conditions, under the weak condition that the error terms follow an autoregressive moving average (ARMA) model, this estimator has the same asymptotic probability distribution as the LS estimator and converges probabilistically to the model coefficients. Finally, a simulation study confirms that the mentioned properties of the new AC-based estimator hold true even in small samples.

DOI: [10.33899/ijoss.v23i1.62136](https://doi.org/10.33899/ijoss.v23i1.62136), ©Authors, 2026, College of Computer Science and Mathematics University of Mosul.

This is an open access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The MLR model refers to a statistical technique as an extension of the simple linear regression model that is used to predict the outcome of one variable (dependent variable) based on the value of two or more variables (independent variables) (Pollins & Cooper, 2022; Mendenhall, 2019). Previous research has shown that two common methods for estimating MLR model coefficients include the LS estimation method and the maximum likelihood (ML) estimation method (Gallo & McKinney, 2025; Kahane, 2025). The LS estimation method is an easy-to-apply estimation method that only considers zero mean and uncorrelated error terms without any distributional assumptions (Piradl et al. 2022; Chen et al. 2021). The estimator obtained with this estimation method is known as the best linear unbiased estimator, according to the Gauss-Markov theorem (Du et al. 2023; Lawrence, 2020). The ML estimation method is a method for estimating the coefficients of the MLR model by adding an assumption to the basic assumptions about the probability distributions of the error terms (Pitblado et al. 2024; Cao et al. 2023). However, when assuming a population with a normal probability distribution, the LS estimator becomes identical to the ML estimator and thus preserves several asymptotic properties of the ML estimator (Magnus, 2022; Spanos, 2019; Osborne, 2006).

This paper introduces an AC-based estimation method for MLR model coefficients. This method uses first-order AC (Bose & Bhattacharjee, 2018; Brockwell, 2009) instead of the sum of squares to minimize or maximize the AC of the error terms. In principle, this method can be used not only for uncorrelated error terms but also for cases where the error terms are autocorrelated. If the AC values are always positive (or negative), the estimator with the minimum (or maximum) AC will be the optimal estimator, with AC close to zero. However, if there is an estimator with zero AC, it

should be the optimal estimator, but in reality, satisfactory results cannot be achieved by obtaining an estimator that is over-adjusted to make the AC zero. From this perspective, this method is not intuitively appealing, but a method for zeroing AC in a different direction can be developed and applied to nonparametric regression (Dudáš, 2025; Kong et al., 2024; Takezawa, 2005).

This paper first introduces an AC-based estimator using matrix notation. Then, it presents two properties of this new estimator. The first property is that this estimator is an unbiased estimator of the coefficients of the MLR model, and the second property is that if the vector of independent variables satisfies certain regular conditions, this estimator has an asymptotic probability distribution similar to the LS estimator and converges in probability to the coefficients of the MLR model, under the weak condition that the error terms follow the ARMA model. Finally, through a simulation study, it is shown that these two properties of the proposed estimator hold true even for small samples.

The content of this paper is organized as follows: Section 2 presents the new AC-based estimator and its properties. Section 3 shows through a simulation study that the properties of the AC-based estimator presented in Section 2 hold true even in small samples. Finally, Section 4 includes discussion and conclusions.

Definition of an AC-Based estimator and presentation of its properties

This section presents the new AC-based estimator and its two main features. Specifically, Section 2.1 introduces this estimator and Section 2.2 demonstrates its unbiasedness and asymptotic equality with the LS estimator.

1.1. Definition of AC-based estimator

This section presents the AC-based estimator using the following definitions and Theorem 1:

Definition 1. The MLR model is considered as follows:

$$Y_l = \theta_0 + \theta_1 x_{l1} + \theta_2 x_{l2} + \dots + \theta_n x_{ln} + \tau_l = \mathbf{x}_l^T \boldsymbol{\theta} + \tau, \quad l = 1, 2, 3, \dots, m, \tag{1}$$

where Y_l , $l = 1, 2, 3, \dots, m$ are dependent random variables, $\mathbf{x}_l^T = (1, x_{l1}, x_{l2}, \dots, x_{ln})$, $l = 1, 2, 3, \dots, m$ are vectors of independent non-random variables, $\boldsymbol{\theta}^T = (\theta_0, \theta_1, \theta_2, \dots, \theta_n)$ is the vector of model coefficients and τ_l , $l = 1, 2, 3, \dots, m$ are error terms, which are independent and identically distributed (i.i.d.) random variables with a common mean equal to 0 and a common variance equal to the unknown non-negative constant σ^2 .

It should be noted that the equation (1) can be written in matrix form as follows:

$$\mathbf{Y} = \mathbf{x}\boldsymbol{\theta} + \boldsymbol{\tau}, \tag{2}$$

where \mathbf{Y} is a random vector with $m \times 1$ elements, \mathbf{x} is a non-random matrix with $m \times (n + 1)$ elements, $\boldsymbol{\theta}$ is a non-random vector with $(n + 1) \times 1$ elements, and $\boldsymbol{\tau}$ is a random vector with $m \times 1$ elements.

Definition 2. The upper shift matrix \mathbf{M} with $m \times m$ elements is considered as follows:

$$\mathbf{M} = (m_{lk}) = \begin{bmatrix} 0 & 1 & 0 & & 0 & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & & 0 & 0 & 0 \\ \vdots & & & \ddots & \vdots & & \\ 0 & 0 & 0 & & 0 & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 1 \\ 0 & 0 & 0 & & 0 & 0 & 0 \end{bmatrix}, \quad \begin{cases} l = 1, 2, 3, \dots, m \\ k = 1, 2, 3, \dots, m' \end{cases} \tag{3}$$

where $m_{lk} = \begin{cases} 1, & \text{if } l + 1 = k \\ 0, & \text{otherwise} \end{cases}$.

Definition 3. Assuming the MLR model $Y_l = \mathbf{x}_l^T \boldsymbol{\theta} + \tau$, $l = 1, 2, 3, \dots, m$ defined in equation (1), the estimator of the model coefficients based on AC, called $\hat{\boldsymbol{\theta}}$, is obtained by setting the partial derivative of the following equation with respect to $\boldsymbol{\theta}$ to zero:

$$K(\boldsymbol{\theta}) = \sum_{l=1}^{m-1} (\tau_l \tau_{l+1}) = \sum_{l=1}^{m-1} (Y_l - \mathbf{x}_l^T \boldsymbol{\theta})(Y_{l+1} - \mathbf{x}_{l+1}^T \boldsymbol{\theta}), \quad (4)$$

where, $K(\boldsymbol{\theta})$ can be expressed in matrix form as follows:

$$K(\boldsymbol{\theta}) = (\mathbf{Y} - \mathbf{x}\boldsymbol{\theta})^T \mathbf{M}(\mathbf{Y} - \mathbf{x}\boldsymbol{\theta}). \quad (5)$$

From the equation (5), an AC-based estimator is obtained based on the following theorem:

Theorem 1. In the MLR model $\mathbf{Y} = \mathbf{x}\boldsymbol{\theta} + \boldsymbol{\tau}$, if the inverse of the matrix $\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}$ exists, then the estimator for the model coefficients based on AC is as follows:

$$\hat{\boldsymbol{\theta}} = [\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}]^{-1} \mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{Y}. \quad (6)$$

Proof: By simple differentiation of $K(\boldsymbol{\theta})$ defined in equation (6), the following relations hold:

$$\frac{\partial K(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = -\mathbf{x}^T \mathbf{M}(\mathbf{Y} - \mathbf{x}\boldsymbol{\theta}) - \mathbf{x}^T \mathbf{M}^T(\mathbf{Y} - \mathbf{x}\boldsymbol{\theta}) = \mathbf{0}$$

$$\Leftrightarrow \mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}\boldsymbol{\theta} = \mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{Y}$$

$$\Leftrightarrow \hat{\boldsymbol{\theta}} = [\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}]^{-1} \mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{Y}. \blacksquare$$

1.2. Characteristics of the AC-Based Estimator

This section first shows, using Theorem 2, that the estimator $\hat{\boldsymbol{\theta}}$ is an unbiased estimator of the coefficient vector $\boldsymbol{\theta}$.

Theorem 2. In the MLR model, $\mathbf{Y} = \mathbf{x}\boldsymbol{\theta} + \boldsymbol{\tau}$, if the inverse of the matrix $\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}$ exists, then $E(\boldsymbol{\tau}) = \mathbf{0}$ and if so, then $E(\hat{\boldsymbol{\theta}}) = \boldsymbol{\theta}$.

Proof: It is easy to see that the following equations hold:

$$E(\boldsymbol{\tau}) = E(\mathbf{Y} - \mathbf{x}\boldsymbol{\theta}) = E(\mathbf{Y}) - \mathbf{x}\boldsymbol{\theta} = \mathbf{x}\boldsymbol{\theta} - \mathbf{x}\boldsymbol{\theta} = \mathbf{0},$$

and so, the conclusion is drawn:

$$E(\hat{\boldsymbol{\theta}}) = [\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}]^{-1} [\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)] E(\mathbf{Y}) = [\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}]^{-1} [\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)] E(\mathbf{x}\boldsymbol{\theta} + \boldsymbol{\tau})$$

$$= [\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}]^{-1} [\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)] [E(\mathbf{x}\boldsymbol{\theta}) + E(\boldsymbol{\tau})]$$

$$= [\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}]^{-1} [\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)] (\mathbf{x}\boldsymbol{\theta} + \mathbf{0}) = [\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}]^{-1} [\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}] \boldsymbol{\theta} = \boldsymbol{\theta}. \blacksquare$$

It is worth noting that if the error terms have a common mean of zero and are uncorrelated, i.e., $E(\boldsymbol{\tau}) = \mathbf{0}$, and their covariance is $Cov(\boldsymbol{\tau}) = \sigma^2 \mathbf{I}$ (where \mathbf{I} is the identity matrix with $m \times m$ elements), then the LS estimator is the best linear unbiased estimator by the Gauss-Markov theorem. Therefore, under the assumptions of typical MLR model analysis, an AC-based estimator has a larger variance than the LS estimator. As such, this paper seeks conditions under which the optimal AC-based estimator is asymptotically identical to the LS estimator. These conditions can be stated as follows:

- (1): The existence of the inverse of the matrices $\mathbf{x}^T \mathbf{x}$ and $\mathbf{x}^T (\mathbf{M} + \mathbf{M}^T) \mathbf{x}$,
- (2): For each $l = 1, 2, 3, \dots, m$, there exists a constant value $\varphi > 0$, such that $|\mathbf{x}_l| \leq \varphi$, where $|\mathbf{x}_l|$ is the notation of L^2 -norm with the dimension defined in the text,
- (3): For each $l = 2, 3, \dots, m$, there is a constant $\gamma > 0$, such that $|\mathbf{x}_l - \mathbf{x}_{l-1}| \leq \frac{\gamma}{m}$,
- (4): $\left(\frac{\mathbf{x}^T \mathbf{x}}{m}\right)^{-1} = \mathbf{O}(1)$, where $\mathbf{O}(1)$ is the notation indicating that the matrix $\left(\frac{\mathbf{x}^T \mathbf{x}}{m}\right)^{-1}$ is bounded,
- (5): Stationary property of elements of the random vector $\boldsymbol{\tau}$, $E(\boldsymbol{\tau}) = 0$, and $\sum_{i=0}^{\infty} |\vartheta(i)| < \infty$, where $\vartheta(i) = cov(\tau_1, \tau_{1+i}) = cov(\tau_m, \tau_{m-i}) = \vartheta(-i)$.

However, it should be noted that the above conditions are strong and can be further modified. The possibility of this adjustment will be discussed in more detail in Section 4. Furthermore, condition (5) generally holds for uncorrelated data and is also valid in the ARMA model (Bisette & Schwartz, 2025).

Now, using conditions (1) to (5), the following lemmas and theorem hold:

Lemma 1. For each $l = 1, 2, 3, \dots, m$, if there is a constant $\delta > 0$, such that $|\mathbf{x}_l| \leq \delta$, and a random vector $\boldsymbol{\tau}$ with multivariate normal probability distribution such that $E(\boldsymbol{\tau}) = 0$, and $\sum_{i=0}^{\infty} |\vartheta(i)| < \infty$, then $\frac{1}{m} \sum_{l=1}^m (\mathbf{x}_l \tau_l) = \mathbf{O}_P(1)$, where $\mathbf{O}_P(1)$ is a random vector that converges to 0 in probability.

Proof: Some simple calculations can lead to the following results:

$$\begin{aligned} \left| cov \left[\frac{1}{m} \sum_{l=1}^m (\mathbf{x}_l \tau_l) \right] \right| &= \left| \frac{1}{m^2} \sum_{l=1}^m \sum_{v=1}^m [\mathbf{x}_l \mathbf{x}_v^T \vartheta(l-v)] \right| \leq \frac{1}{m^2} \sum_{l=1}^m \sum_{v=1}^m [|\mathbf{x}_l \mathbf{x}_v^T| |\vartheta(l-v)|] \\ &\leq \frac{\delta^2}{m^2} \sum_{l=1}^m \sum_{v=1}^m |\vartheta(l-v)| = \frac{\delta^2}{m} \sum_{i=-(m-1)}^{m-1} \left[\left(1 - \frac{|i|}{m}\right) |\vartheta(i)| \right] \leq \frac{\delta^2}{m} \sum_{i=-(m-1)}^{m-1} |\vartheta(i)| \leq \frac{2\delta^2}{m} \sum_{i=0}^{m-1} |\vartheta(i)|, \end{aligned}$$

where the second inequality hold because $|\mathbf{x}_l \mathbf{x}_v^T| = |\mathbf{x}_l| |\mathbf{x}_v^T| \leq \delta^2$, and the last inequality holds because $\vartheta(i) = \vartheta(-i)$. Therefore, if $m \rightarrow \infty$, then $\left| cov \left[\frac{1}{m} \sum_{l=1}^m (\mathbf{x}_l \tau_l) \right] \right| \rightarrow 0$ holds and $\frac{1}{m} \sum_{l=1}^m (\mathbf{x}_l \tau_l)$ converges to 0 in the L^2 -norm, and consequently, $\frac{1}{m} \sum_{l=1}^m (\mathbf{x}_l \tau_l)$ converges to 0 in probability.

Lemma 2. If conditions (1) to (5) hold, the following results are obtained:

- (1): $\mathbf{x}^T (\mathbf{M} + \mathbf{M}^T) \mathbf{x} = 2\mathbf{x}^T \mathbf{x} + \mathbf{O}(1)$,
- (2): $\mathbf{x}^T (\mathbf{M} + \mathbf{M}^T) \boldsymbol{\tau} = 2\mathbf{x}^T \boldsymbol{\tau} + \mathbf{O}_P(1)$,
- (3): $\mathbf{x}^T (\mathbf{M} + \mathbf{M}^T) \mathbf{Y} = 2\mathbf{x}^T \mathbf{Y} + \mathbf{O}_P(1)$.

where $\mathbf{O}_P(1)$ is a probability vector (That is, $\mathbf{O}_P(1)$ is a vector with non-negative elements whose sum is one).

Proof:

(1):

$$\begin{aligned} \mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x} - 2\mathbf{x}^T\mathbf{x} &= \sum_{l=1}^{m-1} (\mathbf{x}_l\mathbf{x}_{l+1}^T) + \sum_{l=1}^{m-1} (\mathbf{x}_{l+1}\mathbf{x}_l^T) - 2\sum_{l=1}^m (\mathbf{x}_l\mathbf{x}_{l+1}^T) \\ &= \sum_{l=1}^{m-1} \mathbf{x}_l (\mathbf{x}_{l+1}^T - \mathbf{x}_l^T) + \sum_{l=1}^{m-1} (\mathbf{x}_{l+1} - \mathbf{x}_l) \mathbf{x}_l^T - 2\mathbf{x}_m\mathbf{x}_m^T = \mathbf{O}(1) + \mathbf{O}(1) + \mathbf{O}(1) \\ &= \mathbf{O}(1), \end{aligned}$$

where the penultimate equation holds because $|\mathbf{x}_l(\mathbf{x}_{l+1}^T - \mathbf{x}_l^T)|$ and $|(\mathbf{x}_{l+1} - \mathbf{x}_l)\mathbf{x}_l^T|$ both have an upper bounds equal to $\frac{\varphi\gamma}{m}$ according to conditions (1) and (2),

(2):

$$\begin{aligned} \mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\boldsymbol{\tau} - 2\mathbf{x}^T\boldsymbol{\tau} &= \sum_{l=1}^{m-1} (\mathbf{x}_l\boldsymbol{\tau}_{l+1}) + \sum_{l=1}^{m-1} (\mathbf{x}_{l+1}\boldsymbol{\tau}_l) - 2\sum_{l=1}^m (\mathbf{x}_l\boldsymbol{\tau}_l) \\ &= \sum_{l=1}^{m-1} (\mathbf{x}_l - \mathbf{x}_{l+1}) \boldsymbol{\tau}_{l+1} + \sum_{l=1}^{m-1} (\mathbf{x}_{l+1} - \mathbf{x}_l) \boldsymbol{\tau}_l - \mathbf{x}_1\boldsymbol{\tau}_1 - \mathbf{x}_m\boldsymbol{\tau}_m = \mathbf{o}_P(1) + \mathbf{o}_P(1) + \mathbf{O}_P(1) \\ &= \mathbf{O}_P(1), \end{aligned}$$

where the penultimate equation holds according to the Lemma 1 because $|\mathbf{x}_l - \mathbf{x}_{l+1}|$ and $|\mathbf{x}_{l+1} - \mathbf{x}_l|$ both have an upper bounds equal to $\frac{\gamma}{m}$ according to condition (2),

(3):

$$\begin{aligned} \mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{Y} &= \mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)(\mathbf{x}\boldsymbol{\theta} + \boldsymbol{\tau}) = \mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}\boldsymbol{\theta} + \mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\boldsymbol{\tau} = 2\mathbf{x}^T\mathbf{x}\boldsymbol{\theta} + \mathbf{O}(1) + 2\mathbf{x}^T\boldsymbol{\tau} + \mathbf{O}_P(1) = \\ &= 2\mathbf{x}^T(\mathbf{x}\boldsymbol{\theta} + \boldsymbol{\tau}) + \mathbf{O}_P(1) = 2\mathbf{x}^T\mathbf{Y} + \mathbf{O}_P(1), \end{aligned}$$

where the penultimate equation holds with results (1) and (2) of Lemma 2.

By combining the previous two lemmas, we obtain the next theorem. This theorem proves that if conditions (1) to (5) hold, then the AC-based estimator has the same asymptotic probability distribution as the LS estimator and also converges in probability to the coefficients of the MLR model.

Theorem 3. If an MLR model of the form $\mathbf{Y} = \mathbf{x}\boldsymbol{\theta} + \boldsymbol{\tau}$ satisfies conditions (1) to (5), then the following hold:

$$(1): \hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_{LS} + \mathbf{O}_P\left(\frac{1}{m}\right),$$

$$(2): \hat{\boldsymbol{\theta}} = \boldsymbol{\theta} + \mathbf{o}_P(1),$$

where $\hat{\boldsymbol{\theta}}_{LS}$ is the LS estimator of the coefficients vector $\boldsymbol{\theta}$, and $\mathbf{O}_P\left(\frac{1}{m}\right)$ is a probability vector for which $\frac{\mathbf{O}_P\left(\frac{1}{m}\right)}{\frac{1}{m}}$ is bounded in probability.

Proof:

(1):

$$\begin{aligned} \hat{\theta} &= [\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}]^{-1}\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{Y} = \left[\frac{\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{x}}{m} \right]^{-1} \frac{\mathbf{x}^T(\mathbf{M} + \mathbf{M}^T)\mathbf{Y}}{m} \\ &= \left[2 \frac{\mathbf{x}^T\mathbf{x}}{m} + \mathbf{O}\left(\frac{1}{m}\right) \right]^{-1} \left[2 \frac{\mathbf{x}^T\mathbf{Y}}{m} + \mathbf{O}_P\left(\frac{1}{m}\right) \right] = \left[\left(2 \frac{\mathbf{x}^T\mathbf{x}}{m} \right)^{-1} + \mathbf{O}\left(\frac{1}{m}\right) \right] \left[2 \frac{\mathbf{x}^T\mathbf{Y}}{m} + \mathbf{O}_P\left(\frac{1}{m}\right) \right] \\ &= (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\mathbf{Y} + \mathbf{O}_P\left(\frac{1}{m}\right) = \hat{\theta}_{LS} + \mathbf{O}_P\left(\frac{1}{m}\right), \end{aligned}$$

(2):

$$\begin{aligned} \hat{\theta} &= \hat{\theta}_{LS} + \mathbf{O}_P\left(\frac{1}{m}\right) = (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\mathbf{Y} + \mathbf{O}_P\left(\frac{1}{m}\right) = (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T(\mathbf{x}\theta + \boldsymbol{\tau}) + \mathbf{O}_P\left(\frac{1}{m}\right) \\ &= (\mathbf{x}^T\mathbf{x})^{-1}(\mathbf{x}^T\mathbf{x})\theta + (\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\boldsymbol{\tau} + \mathbf{O}_P\left(\frac{1}{m}\right) = \theta + \mathbf{o}_P(1) + \mathbf{O}_P\left(\frac{1}{m}\right) = \theta + \mathbf{o}_P(1), \end{aligned}$$

where the penultimate equation holds according to condition (4) and Lemma 1. That is:

$$(\mathbf{x}^T\mathbf{x})^{-1}\mathbf{x}^T\boldsymbol{\tau} = \left(\frac{\mathbf{x}^T\mathbf{x}}{m}\right)^{-1} \frac{\mathbf{x}^T\boldsymbol{\tau}}{m} = \mathbf{o}_P(1).$$

2. A Simulation Study

This section conducts a simulation study to verify the properties of the AC-based estimator theoretically obtained in Section 2 under small samples. In this simulation study, what is considered as the probability distribution of the error terms is that the AC-based estimator, $\hat{\theta}$, and the LS estimator, $\hat{\theta}_{LS}$, have the same probability distribution asymptotically, according to the Theorem 3, and the error terms can be expressed as an $ARMA(r, s)$ model as follows:

$$\tau_z = \sum_{w=1}^r \tau_{z-w}\alpha(w) + \pi_z + \sum_{\omega=1}^s \pi_{z-\omega}\beta(\omega), \quad z \in \mathbb{Z}, \quad (7)$$

where $\{\pi_z\}_{z \in \mathbb{Z}}$ is a sequence of uncorrelated random variables with a common mean of 0.

Furthermore, in this simulation study, four special cases for model defined in the equation (7) are considered: a simple independent sample ($\alpha(1) = 0$ and $\beta(1) = 0$), the $AR(1)$ model only with $\alpha(1) = \frac{1}{2}$ (and $\beta(1) = 0$), the $MA(1)$ model only with $\beta(1) = -\frac{1}{2}$ (and $\alpha(1) = 0$), and the $ARMA(1,1)$ model only with $\alpha(1) = \frac{1}{2}$ and $\beta(1) = -\frac{1}{2}$. The probability distribution of the sequence $\{\pi_z\}_{z \in \mathbb{Z}}$ is assumed to be a normal probability distribution with a common mean of 0. Also, the variance of each $\tau_z, z \in \mathbb{Z}$ is set to 1 to allow comparison between the above four models. Given these assumptions, the MLR model is considered as follows:

$$Y_l = 1 + x_{l2} + x_{l3} + \tau_l, \quad l = 1, 2, 3, \dots, m, \quad (8)$$

$$\text{where } x_{l2} = \frac{l}{m} \text{ and } x_{l3} = \begin{cases} \frac{l}{m}, & \frac{l}{m} \leq \frac{1}{2} \\ 1 - \frac{l}{m}, & \frac{l}{m} > \frac{1}{2} \end{cases}.$$

Therefore, it can be seen that this section uses an MLR model with $\theta_0 = 0$ and $\theta_1 = \theta_2 = \theta_3 = 1$. It is easy to see that the independent variables, including $x_{l1} = 1$ corresponding to the width from the origin of the model, satisfy conditions

(1) and (2). Furthermore, it follows that as m increases, $\left(2\frac{x^T x}{m}\right)^{-1}$ converges to $\begin{pmatrix} 7 & -6 & -12 \\ -6 & 12 & 0 \\ -12 & 0 & 48 \end{pmatrix}$ satisfying the condition (4).

The sample size m is considered to be 100, which is not considered a very large size in time series data analysis. In the MLR model defined in equation (8), 1000 samples are generated and for each sample, $\hat{\theta} = \begin{pmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \\ \hat{\theta}_3 \end{pmatrix}$ and $\hat{\theta}_{LS} = \begin{pmatrix} \hat{\theta}_{LS,1} \\ \hat{\theta}_{LS,2} \\ \hat{\theta}_{LS,3} \end{pmatrix}$ are calculated and the sample means (SMs) and standard errors (SEs) are organized in Tables 1 to 4 as follows:

Table 1. SMs and SEs corresponding to the vectors $\hat{\theta}$ and $\hat{\theta}_{LS}$ based on the special case of a simple independent sample ($\alpha(1) = 0$ and $\beta(1) = 0$) for the model defined in equation (7)

Estimators Model		$\hat{\theta}$			$\hat{\theta}_{LS}$		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_{LS,1}$	$\hat{\theta}_{LS,2}$	$\hat{\theta}_{LS,3}$
A simple independent sample ($\alpha(1) = 0$ and $\beta(1) = 0$)	SMs	1.010	0.981	1.000	1.010	0.981	1.003
	SEs	0.266	0.339	0.691	0.264	0.337	0.683

Table 2. SMs and SEs corresponding to the vectors $\hat{\theta}$ and $\hat{\theta}_{LS}$ based on the special case of the AR(1) model only with $\alpha(1) = \frac{1}{2}$ (and $\beta(1) = 0$) for the model defined in equation (7)

Estimators Model		$\hat{\theta}$			$\hat{\theta}_{LS}$		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_{LS,1}$	$\hat{\theta}_{LS,2}$	$\hat{\theta}_{LS,3}$
AR(1) model with only $\alpha(1) = \frac{1}{2}$ (and $\beta(1) = 0$)	SMs	0.991	0.999	1.023	0.990	1.000	1.023
	SEs	0.464	0.594	1.203	0.457	0.587	1.192

Table 3. SMs and SEs corresponding to the vectors $\hat{\theta}$ and $\hat{\theta}_{LS}$ based on the special case of the MA(1) model only with $\beta(1) = -\frac{1}{2}$ (and $\alpha(1) = 0$) for the model defined in equation (7)

Estimators Model		$\hat{\theta}$			$\hat{\theta}_{LS}$		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_{LS,1}$	$\hat{\theta}_{LS,2}$	$\hat{\theta}_{LS,3}$
MA(1) model with only $\beta(1) = -\frac{1}{2}$ (and $\alpha(1) = 0$)	SMs	1.000	0.996	1.005	1.000	0.996	1.004
	SEs	0.127	0.164	0.310	0.129	0.167	0.313

Table 4. SMs and SEs corresponding to the vectors $\hat{\theta}$ and $\hat{\theta}_{LS}$ based on the special case of the ARMA(1, 1) model only with $\alpha(1) = \frac{1}{2}$ and $\beta(1) = -\frac{1}{2}$ for the model defined in equation (7)

Estimators Model		$\hat{\theta}$			$\hat{\theta}_{LS}$		
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_{LS,1}$	$\hat{\theta}_{LS,2}$	$\hat{\theta}_{LS,3}$
ARMA(1,1) model with only $\alpha(1) = \frac{1}{2}$ and $\beta(1) = -\frac{1}{2}$	SMs	1.000	1.003	1.009	1.000	1.002	1.009
	SEs	0.271	0.337	0.659	0.270	0.333	0.659

Tables 1 to 4 show the values of the SMs and SEs of $\hat{\theta}$ and $\hat{\theta}_{LS}$ based on the four assumed special cases for the model defined in equation (7) for 1000 samples of the size 100. Comparing the SMs of the two estimators, it is found that there is a difference only in their third decimal place. Looking more closely from the perspective of having the unbiased property, it is observed that the values of |estimator - parameter| are all within the 'SE/10' range.

The box plots drawn in figures 1 to 4 summarize the values of $\hat{\theta}$ and $\hat{\theta}_{LS}$ under four specific assumed cases for the model defined in equation (7) for 1000 samples of the size 100 as follows:

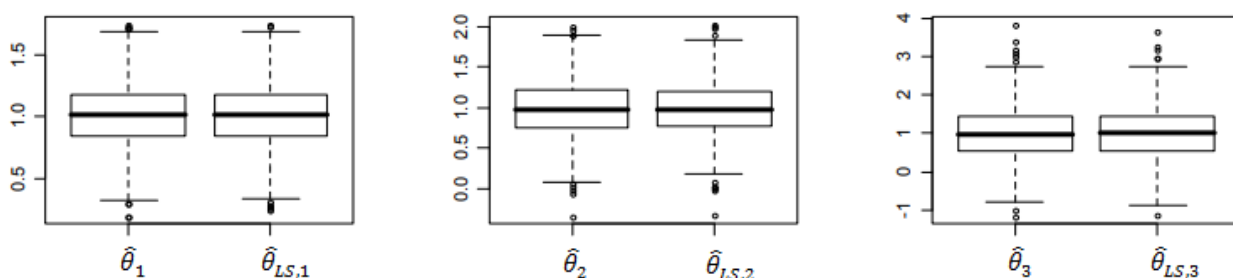


Figure 1. Boxplots corresponding to the vectors $\hat{\theta}$ and $\hat{\theta}_{LS}$ based on the special case of a simple independent sample ($\alpha(1) = 0$ and $\beta(1) = 0$) for the model defined in equation (7)

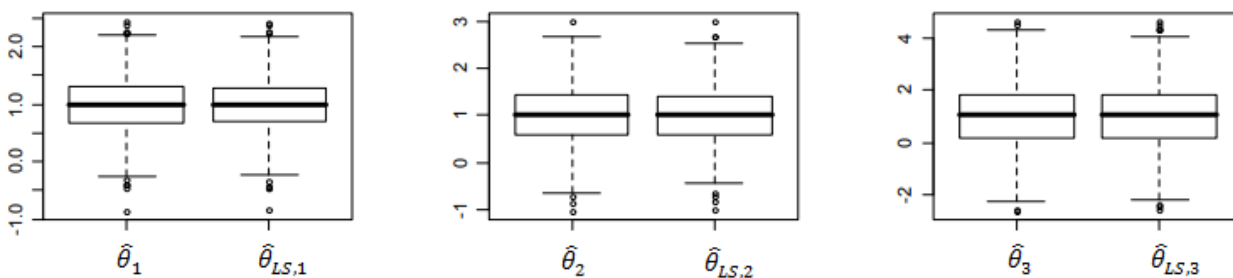


Figure 2. Boxplots corresponding to the vectors $\hat{\theta}$ and $\hat{\theta}_{LS}$ based on the special case of the AR(1) model only with $\alpha(1) = \frac{1}{2}$ (and $\beta(1) = 0$) for the model defined in equation (7)

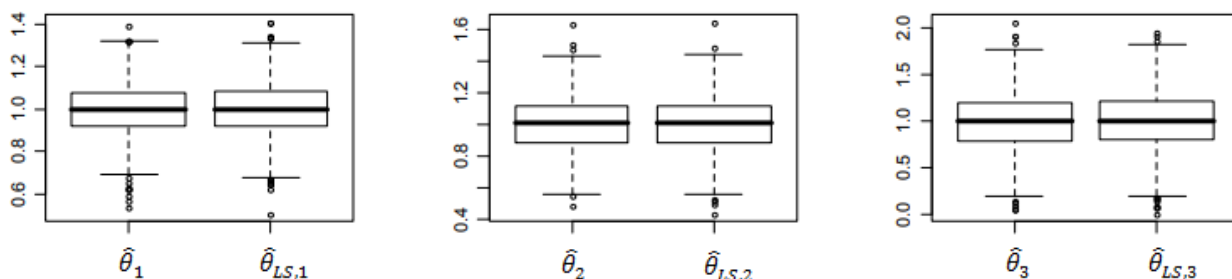


Figure 3. Boxplots corresponding to the vectors $\hat{\theta}$ and $\hat{\theta}_{LS}$ based on the special case of the MA(1) model only with only $\beta(1) = -\frac{1}{2}$ (and $\alpha(1) = 0$) for the model defined in equation (7)

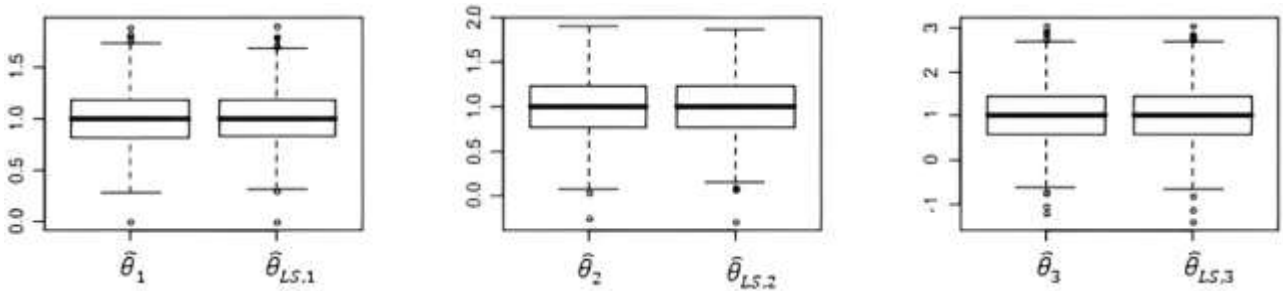


Figure 4. Boxplots corresponding to the vectors $\hat{\theta}$ and $\hat{\theta}_{LS}$ based on the special case of the ARMA(1, 1) model only with $\alpha(1) = \frac{1}{2}$ and $\beta(1) = -\frac{1}{2}$ for the model defined in equation (7)

These box plots are judged based on the SMs and SEs in Tables 1 to 4 and also show that the difference between the AC-based estimator and the LS estimator is very small and the values of both estimators are very close to the true values of the MLR model coefficients.

3. Discussion and Conclusion

This paper presented an AC-based estimator for the coefficients of the MLR model. This new estimator ensures that the derivative of the AC quantity of the error terms with respect to the model coefficients is equal to zero. Then, it was proven that this estimator is an unbiased estimator of the model coefficients. Furthermore, it was shown that if the vector of independent variables satisfies certain regularity conditions, and under the weak condition that the error terms follow an ARMA model, the proposed estimator has a probability distribution similar to the LS estimator asymptotically and converges probabilistically to the model coefficients. Furthermore, the properties of this estimator under small samples were verified through a simulation study in an ARMA model.

Finally, let's examine ways to adjust conditions (1) through (5). Condition (1) is a necessary condition and it is difficult to adjust condition (2) because if condition (2) does not hold, it will be difficult to reach the result of Lemma 1, which plays an important role in this study. Also, if condition (3) or condition (4) does not hold, it is thought that it will not be easy to obtain the results of Lemma 2 or Theorem 3, respectively. Finally, only the condition (5) among the conditions (1) to (5) allows for adjustment. The condition $\sum_{i=0}^{\infty} |\vartheta(i)| < \infty$ in condition (5) was necessary to derive the result of Lemma 1. However, a necessary condition for deriving Lemma 1 is that $\sum_{i=0}^m \frac{|\vartheta(i)|}{m} \rightarrow 0$ holds when $m \rightarrow \infty$. This condition goes one step further and satisfies $\vartheta(m) \rightarrow 0$ when $m \rightarrow \infty$. Therefore, the condition $\sum_{i=0}^{\infty} |\vartheta(i)| < \infty$ can be replaced by a much more adjusted condition $\vartheta(m) = O(1)$. In fact, this condition is also met by models that follow long-term memory processes. Therefore, the range of time series models to which Theorem 3 can be applied is greatly expanded.

Acknowledgments

The authors are very grateful to the Payame Noor University, which helped improve this work's quality.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Ethical Approval

Ethical approval was not required for this study as it did not involve human participants, personal data.

Reference

1. Bisette, V., & Schwartz, A. (2025). Time series analysis in quant finance: Forecasting, volatility modeling, algorithmic trading strategies (kindle ed.). UK: Reactive Publishing.
<https://www.amazon.com.au/Time-Analysis-Quant-Finance-Forecasting/dp/B0FNJ1MJK>
2. Bose, A., & Bhattacharjee, M. (2018). Large covariance and autocovariance matrices (1th ed.). USA: Chapman and Hall/CRC.
<https://www.routledge.com/Large-Covariance-and-Autocovariance-Matrices/BoseBhattacharjee/p/book/9780367734107>
3. Brockwell, P. J. (2009). Autocovariance, WIREs Computational Statistics, 1(2), 187-198.
<https://doi.org/10.1002/wics.20>
4. Cao, Y., Chen, L., Yang, Y., & Chen, J. (2023). Semiparametric maximum likelihood estimation with two-face stratified case-control sampling, Statistica Sinica, 33(3), 2233-2256.
<https://doi.org/10.5705/ss202021.0214>
5. Chen, J., Zhang, X., & Hron, K. (2021). Partial least squares regression with compositional response variables and covariates, Journal of Applied Statistics, 48(16), 3130-3149.
<https://doi.org/10.1080/02664763.2020.1795813>
6. Dudáš, A. (2025). Non-parametric correlation structures and their respective embeddings in predictive analysis, Discover Applied Sciences, 7(5), 1-21.
<https://doi.org/10.1007/s42452-025-07022-0>
7. Du, R., Luo, L., Hudson, L. G., Nozadi, S., & Lewis, J. (2023). An adjusted partial least squares regression framework to utilize additional exposure information in environmental mixture data analysis, Journal of Applied Statistics, 50(8), 1790-1811.
<https://doi.org/10.1080/02664763.2022.2043254>
8. Gallo, M. A., & McKinney, U. O. (2025). Applied multiple regression/Correlation analysis for aviation research (1th ed.). UK: Routledge Taylor & Francis Group.
<https://www.routledge.com/Applied-Multiple-RegressionCorrelation-Analysis-for-Aviation-Research/Gallo-McKinney/p/book/9781032829111>
9. Kahane, L. H. (2025). Regression basics: A student guide to quantitative methods and statistical analysis (3th ed.). UK: Routledge Taylor & Francis Group.
<https://www.routledge.com/Regression-Basics-A-Students-Guide-to-Quantitative-Methods-and-Statistical-Analysis/Kahane/p/book/9781032393186>
10. Kong, D., Zhao, S., & Wang, W. (2024). Nonparametric derivative estimation with bimodal kernels under correlated errors, Computational Statistics, 39(4), 1847-1865.
<https://doi.org/10.1007/s00180-023-01419-4>
11. Lawrence, K. D. (2020). Robust regression (Statistics: A series of textbooks and monographs) (1th ed.). USA: Chapman and Hall/CRC.
<https://www.amazon.com/Robust-Regression-Kenneth-D-Lawrence/dp/0367580187>
12. Magnus, J. R. (2022). Gauss on least-squares and maximum-likelihood estimation, Archive for History of Exact Sciences, 76(4), 425-430.
<https://doi.org/10.1007/s00407-022-00291-w>
13. Mendenhall, W. (2019). A second course in statistics: Regression analysis, (8th ed.). UK: Pearson.
<https://www.amazon.com/Second-Course-Statistics-Regression-Analysis/dp/013516379X>
14. Osborne, M. R. (2006). Least squares methods in maximum likelihood problems, Optimization Methods and Software, 21(6), 943-959.
<https://doi.org/10.1080/10556780600874154>
15. Piradl, S., Shadrokh, A., & Yarmohammadi, M. (2022). A robust estimation method for the linear regression model parameters with correlated error terms and outliers, Journal of Applied Statistics, 49(7), 1663-1676.
<https://doi.org/10.1080/02664763.2021.1881454>
16. Pitblado, J., Poi, B., & Gould, W. (2024). Maximum likelihood estimation with Stata (5th ed.). USA: Stata Press.
<https://www.stata.com/bookstore/maximum-likelihood-estimation-stata/>
17. Pollins, B. M., & Cooper, R. (2022). Advanced regression analysis: An introduction (research method for social scientists) (1th ed.). USA: SAGE publications Ltd.
<https://www.amazon.co.uk/Advanced-Regression-Analysis-Introduction-Scientists/dp/1473912164>

18. Spanos, A. (2019). Probability theory and statistical inference: Empirical modeling with observational data (2th ed.). UK: Cambridge University Press.
<https://www.amazon.com/Probability-Theory-Statistical-Inference-Observational/dp/1107185149>
19. Takezawa, K. (2005). Introduction to nonparametric regression (1th edition). USA: John Wiley & Sons, Inc.
<https://www.amazon.com/Introduction-Nonparametric-Regression-K-Takezawa/dp/0471745839>

مقدر جديد غير متحيز لمعاملات نموذج MLR يعتمد على AC

سجاد بيرعدل

قسم الإحصاء، جامعة بيام نور، طهران، إيران

الخلاصة: يُعدّ نموذج الانحدار الخطي المتعدد (MLR) أحد الأساليب الإحصائية المعروفة في التحليل التنبؤي. وقد اقترحت العديد من الدراسات مقدرات مختلفة لتقدير معاملات هذا النوع من نماذج الانحدار، ومن أشهرها وأكثرها دقةً مقدر المربعات الصغرى العادية (OLS). تُقدّم هذه الورقة البحثية مقدرًا جديدًا لمعاملات نموذج الانحدار الخطي المتعدد (MLR) قائمًا على التغيرات الذاتي (AC). وقد تبين أنّ المقدر القائم على التغيرات الذاتي، على الرغم من أنّه قد لا يبدو بديهيًا للوهلة الأولى، إلا أنّه مقدر غير متحيز لمعاملات النموذج. كما تبين أنّه إذا استوفى متجه المتغيرات المستقلة شروط انتظام معينة، في ظل الشرط الضعيف القائل بأنّ حدود الخطأ تتبع نموذج المتوسط المتحرك التلقائي (ARMA)، فإنّ هذا المقدر له نفس التوزيع الاحتمالي التقاربي لمقبر المربعات الصغرى، ويتقارب احتماليًا مع معاملات النموذج. وأخيرًا، تؤكد دراسة المحاكاة أن الخصائص المذكورة للمقدر الجديد القائم على AC صحيحة حتى في العينات الصغيرة.

الكلمات المفتاحية: نموذج MLR، مقدر غير متحيز، مقدر قائم على AC، مقدر LS، نموذج ARMA.