P- ISSN 1991-8941 E-ISSN 2706-6703
2024,(18), (02):265 – 270
Journal of University of Anbar for Pure Science (JUAPS) Open Access

New Convex Estimator Combining Ridge and Ordinary Least Squares Estimators

Karam Fouad Saray, Mustafa I. N. Alheety

Department of Mathematics, College of education for pure science, University of Anbar, Anbar, Iraq
*E-mail: kar22u2004@uoanbar.edu.iq

ARTICLE INFO
Received: 19 / 05 /2024
Accepted: 01/ 07 /2024
Available online:31/ 12 /2024

10.37652/juaps.2024.149465.1250

Keywords:
Linear Regression Model,
Ridge Estimator,
Multicollinearity,
Convex Estimator

ABSTRACT

In the presence of high correlation between independent variables in the linear regression model, which is known as the multicollinearity problem, the ordinary least squares estimator produces large sample variations. Several estimators have been recommended to overcome this problem. One of these estimators is the estimator of ridge regression, which leads to biased estimated coefficients. However, it has a smaller variance than ordinary least squares estimators, so it may have a smaller mean squared error (MSE). In this study, a new estimator that combines ridge regression and ordinary least squares estimators is proposed. The performance of the recommended estimator is evaluated using the mean square error criterion after obtaining the estimator's properties. The proposed estimator outperforms ordinary least squares and ridge regression estimators in terms of MSE. In addition, the performance of the proposed estimator is studied via simulation and by using a set of real data to demonstrate the theoretical results.

Introduction

The multiple linear regression model is given as follows:

$Y = X\beta + U$. (1)

In this context, Y is a q×1 vector that represents the dependent variable of observations, X is a q×t matrix that represents the independent variable of observations, β is a t×1 vector that represents the unknown parameters that need to be estimated, and U is a q×1 vector that represents normally distributed random variables with a mean of 0 and variance of $\sigma^2 I$ (I is the identity matrix). Ordinary least squares (OLS) is one of the most widely used methods to estimate parameter β of the model in Eq. (1). The OLS of β is given by

$\hat{\beta} = S^{-1}X'Y$, (2)

where $S = X'X$.

In the presence of a multicollinearity issue (MC) among explanatory variables, the OLS estimator becomes unstable and displays undesirable characteristics. Among these characteristics, inflated variance and wide confidence intervals can cause erroneous conclusions and even wrong signs in estimations in extreme cases.

Hoerl and Kennard proposed the ordinary ridge regression (ORR) estimator in 1970 as a way to resolve the MC problem; this technique lowers estimators' excessive variance. Despite its widespread use, the ORR estimator has limitations when it is faced with MC. The ORR estimator has many issues because it is k-dependent. According to Özkale and Kaçiranlar [4], when k approaches infinity, the estimator $\hat{\beta}(k)$ yields zero, which is stable but biased. When k approaches zero, the estimator $\hat{\beta}(k)$ yields $\hat{\beta}_{OLS}$, which is unbiased but stochastic.

Although both estimators are suitable for the given case, a convex combination of both might be helpful [3]. The convex combination's mean squared error (MSE) cannot be higher than the individual MSEs of the two estimators.

This study aims to present a new estimator that combines convex OLS and ORR estimators. MSE and several estimation methods are discussed in Section 2. A new convex estimator that combines OLS and ORR is presented in Section 3. Our results show that the new convex estimator's MSE cannot be higher than the individual MSEs of OLS and ORR. Moreover, an approximation for bias parameter k is calculated. In

*Corresponding author at: Department of Mathematics, College of education for pure science, University of Anbar, Anbar, Iraq
ORCID:https:// https://orcid.org/0000-0000-0000-0000 ,
Tel: +964 7810957909
Email: kar22u2004@uoanbar.edu.iq

Section 4, a numerical example is given to show the theoretical results.

2 Some Estimation Methods

The ORR technique is defined by the bias parameter (k), which is a fixed number. The data matrix's (S) diagonal members are added together. As a result, the estimated parameter variances are reduced, which is desirable, and the values of the diagonal elements of the inverse of S decrease. The estimate is obtained by finding the minimum value of $(Y - X\beta)'(Y - X\beta)$ with respect to $\beta'\beta = c$, where c is a constant. The result is

$(Y - X\beta)'(Y - X\beta) + k(\beta'\beta - c)$, (3)

where k is a Lagrangian multiplier. The ORR estimator is obtained from the solution of Eq. (3) as follows:

$\hat{\beta}(k) = (X'X + kI)^{-1} X'Y, \quad 0 \le k \le 1.$ (4)

In this case, k is selected in accordance with certain fair standards.

This form expresses convex estimator $\hat{\beta}_c$, and it is a product of two estimators.

$\hat{\beta}_c = A\tilde{\beta}_1 + (I - A)\tilde{\beta}_2$, (5)

where A belongs to the set of all $t \times t$ matrices and $\tilde{\beta}_1$ and $\tilde{\beta}_2$ are any two estimators.

2.1 Matrix Mean Square Error Criterion

The MSE matrix, which displays the estimator's proximity to the true parameter value and is given in the following expression, is the standard by which an estimator is judged to be superior.

$$M = M(\tilde{\beta}, \beta) = E(\tilde{\beta} - \beta)(\tilde{\beta} - \beta)'$$

The covariance matrix of $\tilde{\beta}$ is given by

$$cov(\tilde{\beta}) = E\left(\tilde{\beta} - E(\tilde{\beta})\right)\left(\tilde{\beta} - E(\tilde{\beta})\right)'.$$

The difference between $E(\tilde{\beta})$ and β is

$$Bias(\tilde{\beta}, \beta) = E(\tilde{\beta}) - \beta.$$

MSE can be rewritten as follows:

$$M = M(\tilde{\beta}, \beta) = cov(\tilde{\beta}) + \left(Bias(\tilde{\beta}, \beta)\right)\left(Bias(\tilde{\beta}, \beta)\right)'.$$

The scalar MSE (smse) of estimator $\tilde{\beta}$ is given by

$$m = tr\left(M(\tilde{\beta}, \beta)\right) = E(\tilde{\beta} - \beta)'(\tilde{\beta} - \beta),$$

where the symbol $tr\left(M(\tilde{\beta}, \beta)\right)$ stands for the trace of matrix $M(\tilde{\beta}, \beta)$.

3 Proposed Estimator

By combining OLS and ORR, we propose a new version of the convex estimator as follows:

$$\hat{\beta}_{MM} = A\,\hat{\beta}_{OLS} + (I - A)\hat{\beta}(k).$$

We call this new estimator the convex OLS–ORR estimator (COLRR).

Let

$M = M(\hat{\beta}_{MM}, \beta)$, $M_{ij} = E(\tilde{\beta}_i - \beta)(\tilde{\beta}_j - \beta)'$, $i, j = 1,2.$ (6)

According to Odell [5], the MSE of $\hat{\beta}_{MM}$ is minimized when

$A = [M_{22} - M_{21}][M_{11} + M_{22} - M_{12} - M_{21}]^{-1}.$ (7)

Consequently, the minimal value of M is given by

$M_0 = M_{22} - [M_{22} - M_{21}][M_{11} + M_{22} - M_{12} - M_{21}]^{-1}[M_{22} - M_{21}]$ (8)

or

$M_0 = M_{11} - [M_{11} - M_{12}][M_{11} + M_{22} - M_{12} - M_{21}]^{-1}[M_{11} - M_{12}].$ (9)

The following theory shows the advantage of $\hat{\beta}_{MM}$ over OLS and ORR.

Theorem 3.1. In the linear regression model, the MSE of COLRR is smaller than that of OLS and ORR. That is,

$$A = k^2 S_k^{-1}\beta\beta'[\sigma^2 S^{-1}(I - S^2 S_k^{-2})S_k + k^2 S_k^{-1}\beta\beta']^{-1}.$$

Proof: Let $\tilde{\beta}_1 = \hat{\beta}_{OLS}$ and $\tilde{\beta}_2 = \hat{\beta}(k)$. From Eq. (6), we obtain

$$M_{11} = \sigma^2 S^{-1},$$

$$M_{22} = \sigma^2 XX'(X'X + kI)^{-2} + k^2(X'X + kI)^{-1}\beta\beta'(X'X + kI)^{-1}$$
$$= \sigma^2 SS_k^{-2} + k^2 S_k^{-1}\beta\beta' S_k^{-1},$$

where $S = X'X$ and $S_k = X'X + kI$.

$$M_{12} = \sigma^2 XX'(XX' + kI)^{-2} = \sigma^2 SS_k^{-2} = M_{21}$$

$$M_{22} - M_{12} = \sigma^2 SS_k^{-2} + k^2 S_k^{-1}\beta\beta' S_k^{-1} - \sigma^2 SS_k^{-2} = k^2 S_k^{-1}\beta\beta' S_k^{-1}$$

$$M_{11} - M_{12} = \sigma^2 S^{-1} - \sigma^2 SS_k^{-2}$$

We can write $\hat{\beta}_{MM}$'s MSE as

$$M = E(\hat{\beta}_{MM} - \beta)(\hat{\beta}_{MM} - \beta)'$$
$$= ATA' - AW_1' - W_1A' + M_{22},$$

where

$T = M_{11} + M_{22} - M_{12} - M_{21}$ and $W_1 = M_{22} - M_{21}$.

When we set the derivative of M with respect to A to zero, we obtain

$$2AT - W_1 = 0.$$

M is minimized when A is the solution to

$AT = W_1$.

Thus,

$A = W_1 T^{-1}$.

Then,

$A = [M_{22} - M_{21}][M_{11} + M_{22} - M_{12} - M_{21}]^{-1}$.

According to Odell [5], the minimal of $M = M_0$, where

$M_0 = M_{22} - [M_{22} - M_{21}][M_{11} + M_{22} - M_{12} - M_{21}]^{-1}[M_{22} - M_{21}]$, (10)

$= M_{11} - [M_{11} - M_{12}][M_{11} + M_{22} - M_{12} - M_{21}]^{-1}[M_{11} - M_{12}]$. (11)

We now have

$T = M_{11} + M_{22} - M_{12} - M_{21}$

$= \sigma^2 S^{-1} + \sigma^2 S S_k^{-2} + k^2 S_k^{-1} \beta\beta' S_k^{-1} - \sigma^2 S S_k^{-2} - \sigma^2 S S_k^{-2}$

$= \sigma^2 S^{-1} + k^2 S_k^{-1} \beta\beta' S_k^{-1} - \sigma^2 S S_k^{-2}$

$= \sigma^2 S^{-1}(I - S^2 S_k^{-2}) + k^2 S_k^{-1} \beta\beta' S_k^{-1}$

$= [\sigma^2 S^{-1}(I - S^2 S_k^{-2}) S_k + k^2 S_k^{-1} \beta\beta'] S_k^{-1}$.

Thus,

$T^{-1} = [M_{11} + M_{22} - M_{12} - M_{21}]^{-1}$

$= [(\sigma^2 S^{-1}(I - S^2 S_k^{-2}) S_k + k^2 S_k^{-1} \beta\beta') S_k^{-1}]^{-1}$

$= (S_k^{-1})^{-1}[\sigma^2 S^{-1}(I - S^2 S_k^{-2}) S_k + k^2 S_k^{-1} \beta\beta']^{-1}$

$T^{-1} = S_k[\sigma^2 S^{-1}(I - S^2 S_k^{-2}) S_k + k^2 S_k^{-1} \beta\beta']^{-1}$. (12)

Using Eqs. (11) and (12), we obtain

$M_0 - M_{11} = -\sigma^4 S^{-1}(I - S^2 S_k^{-2}) S_k [\sigma^2 S^{-1}(I - S^2 S_k^{-2}) S_k + k^2 S_k^{-1} \beta\beta']^{-1} S^{-1}[I - S^2 S_k^{-2}]$. (13)

Given that

$I - S^2 S_k^{-2} > 0$, $k > 0$, $S_k^{-1} > 0$, $S_k > 0$ and $S^{-1} > 0$, then $M_0 - M_{11} < 0$, where 0 symbolizes a zero matrix. Thus, $M_0 < M_{11}$. Using Eqs. (10) and (12), we obtain

$M_0 - M_{22} = -k^4 S_k^{-1} \beta\beta' [\sigma^2 S^{-1}(I - S^2 S_k^{-2}) S_k + k^2 S_k^{-1} \beta\beta']^{-1} S_k^{-1} \beta\beta' S_k^{-1}$. (14)

Given that

$I - S^2 S_k^{-2} > 0$, $k > 0$, $S_k^{-1} > 0$, $S_k > 0$ and $S^{-1} > 0$, then $M_0 - M_{22} < 0$. Thus, $M_0 < M_{22}$. The proof is completed.

Considering that matrix A in the COLRR estimator depends on unknown parameters β and σ^2, for practicality, we replace β and σ^2 with unbiased estimators $\hat{\beta}$ and $\hat{\sigma}^2 = \frac{1}{n-p}(Y - X\hat{\beta})'(Y - X\hat{\beta})$, respectively. Therefore, the COLRR estimator takes the form

$\hat{\beta}_{MM} = \hat{A}\,\hat{\beta}_{OLS} + (I - \hat{A})\hat{\beta}(k)$,

where $\hat{A} = k^2 S_k^{-1}\hat{\beta}\hat{\beta}'[\hat{\sigma}^2 S^{-1}(I - S^2 S_k^{-2}) S_k + k^2 S_k^{-1}\hat{\beta}\hat{\beta}']^{-1}$.

3.2 Estimation of Bias Parameter k

Bias parameter k must be estimated before the new estimator is implemented. Reducing the MM estimator's smse yields the best k value.

With Eq. (13), the MSE of $\hat{\beta}_{MM}$ can be written as

$M_0 = \sigma^2 S^{-1} - \sigma^4 S^{-1}(I - S^2 S_k^{-2}) S_k [\sigma^2 S^{-1}(I - S^2 S_k^{-2}) S_k + k^2 S_k^{-1} \beta\beta']^{-1} S^{-1}[I - S^2 S_k^{-2}]$.

If an orthogonal matrix D exists with the property $D' X' X D = \Lambda = diag(\lambda_1, \ldots, \lambda_t)$, where Λ is a t×t diagonal matrix and λ_i refers to the eigenvalues of $X'X$, then the regression model in Eq. (1) may be presented in a canonical form by using this form. The considered paradigm is

$Y = T\alpha + \epsilon$,

where $T = XD$, $T'T = \Lambda$, and $\alpha = D'\beta$. The OLS, ORR, and MM estimators in this canonical model are considered as follows:

$\hat{\alpha} = \Lambda^{-1} T' Y$,

$\hat{\alpha}(k) = (\Lambda + kI)^{-1} T' Y$,

$\hat{\alpha}_{MM} = A\hat{\alpha} + (I - A)\,\hat{\alpha}(k)$.

The MSE of the MM estimator is given as

$MSE(\hat{\alpha}_{MM}) = \sigma^2 \Lambda^{-1} - \sigma^4 \Lambda^{-1}(I - \Lambda^2 S_1^{-2}) S_1 [\sigma^2 \Lambda^{-1}(I - \Lambda^2 S_1^{-2}) S_1 + k^2 S_1^{-1}\alpha\alpha']^{-1} \Lambda^{-1}[I - \Lambda^2 S_1^{-2}]$,

where $S_1 = \Lambda + kI$.

The smse of MM is given by

$$smse\ (\hat{\alpha}_{MM}) = tr\big(MSE(\hat{\alpha}_{MM})\big)$$
$$= \sum_{i=1}^{P}\left(\frac{\sigma^2}{\lambda_i} - \frac{\sigma^4}{\lambda_i}\left(1-\frac{\lambda_i^2}{(\lambda_i+k)^2}\right)(\lambda_i + k)\left[\frac{\sigma^2}{\lambda_i}\left(1-\frac{\lambda_i^2}{(\lambda_i+k)^2}\right)(\lambda_i+k) + \frac{k^2\alpha^2}{\lambda_i+k}\right]^{-1}\frac{1}{\lambda_i}\left(1-\frac{\lambda_i^2}{(\lambda_i+k)^2}\right)\right).$$

By differentiating $smse\ (\hat{\alpha}_{MM})$ with respect to k and setting $(\partial\ smse(\hat{\alpha}_{MM}/\partial\ k) = 0$, we obtain

$$k_{1,2} = \frac{\sigma^2 \pm \sigma\sqrt{\sigma^2 + 8\alpha^2\lambda_i}}{\lambda\alpha^2}. \quad (15)$$

The choice of k that works best in Eq. (15) is affected by unknown parameters σ^2 and α. Consequently, we substitute their unbiased estimates for the two estimators. Then, we derive

$$\hat{k}_i = \frac{\hat{\sigma}^2 \pm \hat{\sigma}\sqrt{\hat{\sigma}^2 + 8\,\hat{\alpha}_i{}^2\lambda_i}}{2\hat{\alpha}_i{}^2}. \quad (16)$$

From Eq. (16), we propose the following estimators for k.

$$\hat{k}_{MM} = \frac{\min(\hat{k}_i) * \sum_{i=1}^{p}\hat{k}_i}{\hat{\sigma}^2} \quad (17)$$

$$\hat{k}_{HMM} = \frac{\min(\hat{k}_i)}{\hat{\sigma}^2} \quad (18)$$

$$\hat{k}_{SMM} = \frac{n*p*\hat{\sigma}^2}{\min(\hat{k}_i)} \quad (19)$$

Many researchers have proposed different estimates for k. We present some of them as follows:

- Hoerl and Kennard [2] suggested selecting the value of k by using the formula

$$\hat{k}_{Hk} = \frac{\hat{\sigma}^2}{\hat{\beta}^2{}_{\max OLS}}, \quad (20)$$

where $\hat{\beta}^2{}_{\max OLS}$ is the maximum element of $\hat{\beta}_{OLS}$.

- Hoerl et al. [7] presented

$$\hat{k}_{HKB} = \frac{t\,\hat{\sigma}^2}{\hat{\beta}'_{OLS}\,\hat{\beta}_{OLS}}, \quad (21)$$

where $\hat{k}_{HKB}$ is the harmonic-mean version of the biasing parameter for the ORR estimator and t is the number of explanatory variables.

- Lawless and Wang [8] suggested the following k:

$$\hat{k}_{LW} = \frac{t\,\hat{\sigma}^2}{\hat{\beta}'_{OLS}X'X\hat{\beta}_{OLS}}. \quad (22)$$

- Hocking, Speed, and Lynn [13] suggested the following k (denoted here as $\hat{k}_{HSL}$):

$$\hat{k}_{HSL} = \hat{\sigma}^2\frac{\sum_{i=1}^{p}(\lambda_i\hat{\alpha}_i)^2}{\sum_{i=1}^{p}(\lambda_i\hat{\alpha}_i{}^2)^2}.$$

Other researchers, such as Alkhamisi et al. [1], Mohammad and Naif [9], and Lattef and Alheety [6], have introduced estimates for k.

Numerical Example

To verify our theoretical claims, we use a dataset on Portland cement that was originally published by Woods et al. [10] and investigated extensively by researchers, such as Alheety and Kibria [12] and Alheety and Gore [11]. The data in this dataset were derived from an experiment that examined how different kinds of Portland cement solidified and strengthened under heat. This study assesses the correlation between this heat and the ratios of four cement-making clinkers' constituents. The dependent variable is Y, which stands for the amount of heat produced in calories per gram of cement. The quantities of the compounds tricalcium aluminate, tricalcium silicate, tetracelcium aluminate femrite, and dicalcium silicate (X1, X2, X3, and X4, respectively) serve as the independent variables. The intercept item can be found in the model.

We examine how OLS, ORR, and MM estimators construct their MSE matrices and compare their traces. OLS and ORR provide traces of their respective MSE matrices, which are

$$smse(\hat{\alpha}) = tr\big(MSE(\hat{\alpha}_R)\big) = \sum_{i=1}^{P}\frac{\sigma^2}{\lambda_i}, \quad (23)$$

$$smse(\hat{\alpha}_R) = tr\big(MSE(\hat{\alpha}_R)\big) = \sum_{i=1}^{P}\frac{\lambda_i\sigma^2 + k^2\alpha_i{}^2}{(\lambda_i+k)^2}. \quad (24)$$

Matrix X' X has five eigenvalues, namely, $\lambda_1 = 211.367, \lambda_2 = 77.236, \lambda_3 = 28.459, \lambda_4 = 10.267$, and $\lambda_5 = 0.0349$.

X can be classified as ill-conditioned because its condition number is 6,056.37, which is calculated as K = $\lambda_{max}/\lambda_{min}$. For matrix X' X to take the shape of a correlation matrix, data standardization is suggested by the authors. The regression coefficients may therefore be presented as equivalent numerical units, which is an advantage of data standardization [11].

The four eigenvalues of correlation matrix X' X are 2.2357, 1.57606, 0.18661, and 0.00162, and the estimated value of σ^2 is 0.00196.

Table 1. Results of MSE for the estimators and different estimated ridge parameters

k	OLS	ORR	MM
$\hat{k}_{HKB}$	1.21862916	0.93000998	0.29273114
$\hat{k}_{HKB}$	1.21862916	0.49321202	0.15819601
$\hat{k}_{LW}$	1.21862916	0.79922137	0.26079447
$\hat{k}_{HSL}$	1.21862916	0.93000998	0.29273114
$\hat{k}_{MM}$	1.21862916	0.66179729	0.0000272
$\hat{k}_{HMM}$	1.21862916	0.31844794	0.00121719
$\hat{k}_{SMM}$	1.21862916	0.51942667	0.00039823

The numerical findings match the theoretical results, as shown in Table 1. The MM estimator has a smaller smse compared with the OLS and ORR estimators when used separately. Compared with other methods, the suggested k estimators show superior performance. In ORR and MM estimators, all ridge estimators outperform the OLS estimator.

Conclusions

Our theoretical and numerical findings show that the newly proposed MM estimator outperforms OLS and ORR estimators in every scenario. The MSE of any ridge estimator is lower than that of the OLS estimator. In terms of generating small MSE, the newly proposed k estimators surpass all existing estimators, so they may be recommended to practitioners.

Acknowledgments

We dedicate this article to those who lost their lives in Gaza and Palestine.

Conflict of Interest

The authors declare that they have no conflicts of interest.

References

[1] Alkhamisi, M., Khalaf, G., & Shukur, G. (2006). Some modifications for choosing ridge parameters. *Communications in Statistics – Theory and Methods*,35(11), 2005-2020.

[2] Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55-67.

[3] Kaçıranlar, S., Dawoud, I., & Karapınar, D. (2021). Linear Convex Combination Estimators and Comparisons. *Adıyaman University Journal of Science*, *11*(2), 302-320.

[4] Özkale, M. R., & Kaciranlar, S. (2007). The restricted and unrestricted two-parameter estimators. *Communications in Statistics—Theory and Methods*, *36*(15), 2707-2725.

[5] Odell, P. L., Dorsett, D., Young, D., & Igwe, J. (1989). Estimator models for combining vector estimators. *Mathematical and Computer Modeling*, *12*(12), 1627-1642.

[6] Lattef, M. N., & ALheety, M. I. (2020). Study of some kinds of ridge regression estimators in linear regression model. *Tik. J. of Pure Sci.*, *25*(5), 130-142.

[7] Hoerl, A. E., Kannard, R. W., & Baldwin, K. F. (1975). Ridge regression: some simulations. *Communications in Statistics-Theory and Methods*, *4*(2), 105-123.

[8] JF, L., & P, W. (1976). A simulation study of ridge and other regression estimators. *Communications in Statistics-theory and Methods*, *5*(4), 307-323.

[9] Mohammad, B. A., & Naif, M. I. (2021). A simulation study of some restricted estimators in restricted linear regression model. *Tikrit Journal of Pure Science*, *26*(3), 89-101.

[10] Woods, H., Steinour, H. H., & Starke, H. R. (1932). Effect of composition of Portland cement on heat evolved during hardening. *Industrial & Engineering Chemistry*, *24*(11), 1207-1214.

[11] Alheety, M. I., & Gore, S. D. (2008). A new estimator in multiple linear regression model. *Model Assisted Statistics and Applications*, *3*(3), 187-200.

[12] Alheety, M. I., & Kibria, B. G. (2009). On the Liu and almost unbiased Liu estimators in the presence of multicollinearity with heteroscedastic or correlated errors. *Surveys in Mathematics and its Applications*, *4*, 155-167.

[13] Hocking, R. R., Speed, F. M., & Lynn, M. J. (1976). A class of biased estimators in linear regression. *Technometrics*, 18(4),55-67.

مقدر محدب جديد يجمع بين مقدرات الحرف والمربعات الصغرى العادية

كرم فؤاد سراي، مصطفى اسماعيل نايف

قسم الرياضيات، كلية التربية للعلوم الصرفة، جامعة الأنبار، الأنبار، العراق
email: kar22u2004@uoanbar.edu.iq

الخلاصة:

في ظل وجود علاقة خطية متعددة بين المتغيرات المستقلة في نموذج الانحدار الخطي والتي تعرف بمشكلة التداخل الخطي تؤدي إلى ان مقدرات المربعات الصغرى الاعتيادية تنتج تباينات كبيرة في العينة. وللتغلب على هذه المشكلة تمت التوصية بالعديد من المقدرات، من احدى المقدرات التي تم اقتراحها للتقليل من تأثير مشكلة التداخل الخطي هو مقدر انحدار الحرف الذي يؤدي الى معاملات تقديرية متحيزة ولكنها تمتلك تباين أصغر من مقدرات المربعات الصغرى الاعتيادية مما قد يكون لها متوسط مربعات خطأ أصغر (MSE). نحن نقدم مقدراً جديداً في هذه الدراسة يدمج مقدرات انحدار الحرف والمربعات الصغرى الاعتيادية. يتم تقييم اداء المقدر الموصى به باستخدام معيار متوسط مربعات الخطأ بعد الحصول على خصائصه. وقد تبين ان هذا المقدر يتفوق على كل من مقدرات المربعات الصغرى الاعتيادية ومقدرات انحدار الحرف من حيث متوسط مربعات الخطأ. بالإضافة إلى ذلك تم دراسة اداء هذا المقدر المقترح عن طريق استخدام دراسة المحاكاة ومجموعة من البيانات الواقعية لإثبات النتائج النظرية.

الكلمات المفتاحية: نموذج الانحدار الخطي، مقدر الحرف، مشكلة التداخل الخطي، المقدر المحدب.