Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v14n1p254

Generalized ridge estimator shrinkage estimation based on Harris hawks optimization algorithm

By Absulazeez, Algamal

Published: 20 May 2021

Electronic Journal of Applied Statistical Analysis
Vol. 14, Issue 01, May 2021, 254-265
DOI: 10.1285/i20705948v14n1p254

Generalized ridge estimator shrinkage estimation based on Harris hawks optimization algorithm

Qamar Abdulkareem Abdulazeez and Zakariya Yahya Algamal*

Department of Statistics and Informatics, University of Mosul, Mosul, Iraq

Published: 20 May 2021

It is well-known that in the presence of multicollinearity, the ridge estimator is an alternative to the ordinary least square (OLS) estimator. Generalized ridge estimator (GRE) is an generalization of the ridge estimator. However, the efficiency of GRE depends on appropriately choosing the shrinkage parameter matrix which is involved in the GRE. In this paper, a Harris hawks optimization algorithm, which is a metaheuristic continuous algorithm, is proposed to estimate the shrinkage parameter matrix. The simulation study and real application results show the superior performance of the proposed method in terms of prediction error.

keywords: Multicollinearity, shrinkage parameter, generalized ridge estimator, particle swarm optimization.

1 Introduction

Regression modeling is a widely applied strategy for studying several real data problems. In linear regression model, the response variable is considered as a continuous and reasonably assumed to follow normal distribution. In linear regression models, it is assumed that the correlations among the explanatory variables are not high (Alkhamisi and Shukur, 2007; Asar et al., 2014; Dorugade, 2014; Algamal, 2018b; Alobaidi et al., 2021; Rashad and Algamal, 2019; Shamany et al., 2019; Algamal, 2018a; Alkhateeb and Algamal, 2020; Algamal and Asar, 2020; Yahya Algamal, 2019; Algamal and Alanaz, 2018; Algamal, 2018c, 2020; Al-Taweel and Algamal, 2020). However, this assumption is

*Corresponding author: zakariya.algamal@uomosul.edu.iq

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

not always hold in practice. In linear regression model, the ordinary least squares (OLS) estimator is the best estimator among all linear and unbiased estimators. However, under multicollinearity, OLS estimator becomes unhelpful due to their large variance.

The ridge estimator (RE) (Hoerl and Kennard, 1970) has been consistently demonstrated to be an attractive and alternative to the OLS, when the multicollinearity exists. RE can shrink all the regression coefficients toward zero to reduce the large variance (Batah et al., 2008). Generalized ridge estimator (GRE) has also been considered as a generalization of the RE. The performance of the GRE is fully depending on the values of the shrinkage parameter matrix. Accordingly, appropriate choosing of the shrinkage parameter matrix is an important part of applying GRE. Numerous approaches are available for estimating the shrinkage parameter in the literature (Asar et al., 2014; Batah et al., 2008; Hameed et al., 2017).

In recent years, numerous natural-inspired algorithms have been successfully introduced and applied as random search strategies for solving a number of optimization problems. Harris hawks optimization algorithm is a comparatively recent population-based algorithm that is inspired by swarm.

In this paper, the Harris hawks optimization algorithm is proposed to estimate the values of the shrinkage parameter matrix in GRE. Our proposed approach will efficiently help to find the best values with high prediction accuracy. The superiority of our proposed approach in different simulated examples and a real data application is proved.

2 Generalized ridge estimator

Suppose that we have a data set $\{(y_i, x_i)\}_{i=1}^{n}$ where $y_i \in \mathbb{R}$ is a response variable and $x_i = (x_{i1}, x_{i2}, ..., x_{ip}) \in \mathbb{R}^p$ represents a p-dimensional explanatory variable vector. Without loss of generality, it is assumed that the response variable is centered and the explanatory variables are standardized.

Consider the following linear regression model,

$$y = X\beta + \varepsilon, \tag{1}$$

where y is an $n \times 1$ vector of observations of the response variable, $X = (x_1, ..., x_p)$ is an $n \times p$ known design matrix of explanatory variables, $\beta = (\beta_1, ..., \beta_p)$ is a $p \times 1$ vector of unknown regression coefficients, and ε is an $n \times 1$ vector of random errors with mean 0 and variance σ^2. Using OLS method, the parameter estimation of Eq. (1) is given by

$$\hat{\beta}_{OLS} = (X^T X)^{-1} X^T y. \tag{2}$$

OLS estimator is unbiased and it has minimum variance among all linear unbiased estimators. However, in the presence of multicollinearity, the $X^T X$ matrix is nearly singular that makes OLS estimator unstable due to their large variance. To reduce the effects of the multicollinearity, RE , which is the most commonly used method, adds a positive shrinkage parameter, k, to the main diagonal of the $X^T X$ matrix. The RE is defined as

$$\hat{\beta}_{RE} = (X^T X + kI)^{-1} X^T y, \tag{3}$$

where I is the identity matrix with dimension $p \times p$. The estimator $\hat{\beta}_{RE}$is biased but more stable and has less mean square error. The shrinkage parameter, k, controls the shrinkage of β toward zero. The OLS estimator can be considered as a special estimator from the RE with $k = 0$. For larger value of k, the RE yields greater shrinkage approaching zero (Yang and Emura, 2017). Rewriting Eq.(1) as (Alkhamisi and Shukur, 2007)

$$y = Z\alpha + \varepsilon \tag{4}$$

where $Z = XW$,where $W$is a matrix $p \times p$ So that $Z'Z = W'X'XW$will implies $Z'Z = \Lambda = diag(\lambda_1, \lambda_1, ..., \lambda_p)$where Λ is a diagonal matrix with the Eigen values of $X'X$and $\alpha = W'\beta$, then OLS estimator of α is given by:

$$\hat{\alpha}_{LS} = \Lambda^{-1} Z'y \tag{5}$$

Therefore the OLS of β is

$$\hat{\beta} = W\hat{\alpha}_{LS} \tag{6}$$

The RE is given by (Hoerl & Kennard, 1970)

$$\hat{\beta}_{RE} = (\Lambda + kI)^{-1} Z'y = A^{-1} Z'y \tag{7}$$

where $A = \Lambda + kI$.

Relating to Eq. (3) and Eq.(5), the mean square error(MSE) is

$$MSE(\hat{\beta}_{RR}) = \hat{\sigma}^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k)^2} + k^2 \sum_{i=1}^{p} \frac{\alpha_i^2}{(\lambda_i + k)^2} \tag{8}$$

The GRE is suggested by Hoerl and Kennard (1970) to generalize the ridge estimator. The difference between RE and GRE is there are ith values of k, such that

$$\hat{\beta}_{GRE} = (\Lambda + K)^{-1} Z'y = A^{-1} Z'y \tag{9}$$

where $K = diag(k_1, k_2, ..., k_p)$. The MSE, which is less than when using the RE and OLS, is

$$MSE(\hat{\beta}_{GRE}) = \hat{\sigma}^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k_i)^2} + k^2 \sum_{i=1}^{p} \frac{\alpha_i^2}{(\lambda_i + k_i)^2} \tag{10}$$

where $\hat{\sigma}^2 = \frac{y^T y - \hat{\beta}_{OLS} Zy}{n-p-1}$.

Since the ridge parameter is the key to reduce the multicollinearity, there are multi ways to determine this value, the researcher suggest several ways to choose the optimal k such as (Hocking et al., 1976; Nomura, 1988; Troskie and Chalton, 2019; Firinguetti, 1999; Asar et al., 2014; Bhat and Raju, 2017; Alheety and Kibria, 2014; Dorugade and Kashid, 2010; Bhat, 2016). These methods can be defined as:

$$\hat{k}_{i(HK)} = \frac{\hat{\sigma}^2}{\hat{\beta}_i^2} \tag{11}$$

$$\hat{k}_{i(N)} = \frac{\hat{\sigma}^2}{\hat{\beta}_i^2}\left\{1 + \left[1 + k_i(\hat{\beta}_i^2/\hat{\sigma}^2)^{1/2}\right]\right\} \tag{12}$$

$$\hat{k}_{i(TC)} = \frac{k_i\hat{\sigma}^2}{k_i\hat{\beta}_i^2 + \hat{\sigma}^2} \tag{13}$$

$$\hat{k}_{i(F)} = \frac{k_i\hat{\sigma}^2}{k_i\hat{\beta}_i^2 + (n-p)\hat{\sigma}^2} \tag{14}$$

$$\hat{k}_{i(HSL)} = \hat{\sigma}^2\frac{\sum_{i=1}^{p}(k_i\hat{\beta}_i^2)^2}{(\sum_{i=1}^{p}(k_i\hat{\beta}_i^2))^2} \tag{15}$$

$$\hat{k}_{i(AH)} = \hat{\sigma}^2\frac{\sum_{i=1}^{p}(k_i\hat{\beta}_i^2)^2}{(\sum_{i=1}^{p}(k_i\hat{\beta}_i^2))^2} + \frac{1}{\lambda_{\max}} \tag{16}$$

$$\hat{k}_{i(D)} = \frac{\hat{\sigma}^2}{k_{\max}\hat{\beta}_i^2} \tag{17}$$

$$\hat{k}_{i(SB)} = \frac{k_i\hat{\sigma}^2}{k_i\hat{\beta}_i^2 + \hat{\sigma}^2} + \frac{1}{k_{\max}} \tag{18}$$

$$\hat{k}_{i(DK)} = Max\left(o, \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}} - \frac{1}{n(VIF_j)_{Max}}\right) \tag{19}$$

$$\hat{k}_{i(SV1)} = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}} + \frac{1}{k_{Max}\hat{\beta}'\hat{\beta}} \tag{20}$$

$$\hat{k}_{i(SV2)} = \frac{p\hat{\sigma}^2}{\hat{\beta}'\hat{\beta}} + \frac{1}{2\left(\sqrt{k_{Max}/k_{Min}}\right)^2} \tag{21}$$

$$\hat{k}_{i(M)} = \frac{1}{\frac{k_{Max}\hat{\beta}_i^2}{(n-p)\hat{\sigma}^2 + k_{Max}\hat{\beta}_i^2}} \tag{22}$$

$$\hat{k}_{i(AS)} = \frac{\hat{\sigma}^2}{\hat{\beta}_i^2} + \frac{1}{k_i} \tag{23}$$

3 The proposed method

The efficiency of ridge regression model strongly depends on appropriately choosing the shrinkage parameter. A choice of shrinkage parameter that is too small leads to overfitting the GRE, while shrinkage parameter that is too large shrinks β by too much, making a bias-variance tradeoff.

Harris hawks optimization algorithm (HHOA), which was introduced by Heidari et al. (2019), is designed based on simulation of the behaviors of Harris Hawks during the process of search and catching the rabbit in natural space. The optimization process of the HHOA is explained by three phases to find the optimal solution for any given problem. These phases are exploration, transition from exploration to exploitation, and exploitation.

3.1 Exploration phase

The exploration phase mimics the process where a Harris hawk is no able to properly track the prey. When it occurs, the hawks take a break to monitor and identify new preys. In the HHOA, the candidate solutions are the hawks and the best so far solution at each step is the prey. The hawks then perch randomly in a different location and wait for a prey using two operators that are selected based on a probability q.

Mathematically, this process is modeled as

$$x^{(t+1)} = \begin{cases} x^t_{rand} - r_1|x^t_{rand} - 2r_2x^t & q \geq 0.5 \\ (x^t_{prey} - x^t_m) - r_3(L_b + r_4(U_b - L_b)) & q < 0.5, \end{cases} \tag{24}$$

where $x^{(??)}$ is the position vector of hawks in the next iteration, x^t_{prey} represents the position of intended rabbit, and x^t_{rand} is the position of a hawk which is chosen randomly from current team. r_1, r_2, r_3 and r_4 are random numbers. L_b and $U_b$are the upper and lower bounds of search space. x^t_m is the average position of the current population of hawks which is calculated by the following equation

$$x^t_m = \frac{1}{nh}\sum_{i=1}^{nh} x^t_i, \tag{25}$$

where x^t_i is the position of each hawk in team and nh indicates the total number of team members.

3.2 Transition phase

According to the energy level of the prey (escape energy), E, the HHO algorithm goes from the exploration phase to the exploitation phase. The energy reduction of the prey is defined as

$$E = 2E_0(1 - \frac{t}{t_{\max}}), \tag{26}$$

where $t_{\max}$ indicates the maximum number of iterations and $E_0$is the initial energy which is randomly changing inside $(-1, 1)$at each iteration. This value is used to indicate either the physically flagging of the prey for $-1 \leq E_0 < 0$ or its strengthening $0 \leq E_0 < 1$. Furthermore, in the case $|E| \geq 1$ then HHOA will explore the search space otherwise, HHOA will change its status to the exploitation phase.

3.3 Exploitation phase

During the exploitation phase, the $|E|$is considered to choose the type of besiege to catch the prey. Accordingly, a soft one is taken when $|E| \geq 0.5$, and the hard one is taken when $|E| < 0.5$. This process is stimulated by the following two strategies: Soft besiege and hard besiege.

$$x^{(t+1)} = \Delta x^t - E\left|J x^t_{prey} - x^t\right|, \tag{27}$$

where $\Delta x^t = x^t_{prey} - x^t$, $J = 2(1 - r_5)$, which is standing for jump severity of the prey in the stage of escaping, and r_5 is a random number in the range [0,1].

On the other hand, in hard besiege strategy, the $r \geq 0.5$and $|E| < 0.5$, which means that the prey is tired and does not have sufficient energy to escape. The updated position of the Harris' hawk is defined as

$$x^{(t+1)} = x^t_{prey} - E \left|\Delta x^t\right| \ , \tag{28}$$

In case of $r < 0.5$and $|E| \geq 0.5$, which is called soft besiege with progressive rapid dives, the Harris' hawk progressively selects the best possible dive to catch the prey competitively. Then, the new position of the hawk is mathematically modeled as

$$\Upsilon = x^t_{prey} - E \left|J\, x^t_{prey} - x^t\right| . \tag{29}$$

The Harris' hawk can dive by

$$\mathrm{Z} = \Upsilon + S \times \mathrm{L}evy(\mathrm{D}), \tag{30}$$

In this paper, a HHOA algorithm is proposed to determine the shrinkage parameter matrix. The proposed method will efficiently help to reduce the MSE. The parameter configurations for our proposed method are presented as follows.

1. The positions of each hawks are randomly determined. The position of a hawk represents the shrinkage parameters, k_i. Here the dimension of each hawk is the number of explanatory variables. The initial positions of the hawks are generated from a uniform distribution within the range [0,100].
2. The fitness function is the MSE
3. The positions are updated using Eq. (27).
4. Steps 3 and 4 are repeated until a $t_{\max}$ is reached.

4 Monte Carlo simulation results

In this section, a comprehensive simulation study was conducted to evaluate the performance of the proposed method. Following Asar et al. (2014), the explanatory variables with different degree of multicollinearity are generated by

$$x_{ij} = (1 - \rho^2)^{1l2} w_{ij} + \rho w_{ip}, \ \ i = 1, 2, ..., n, \quad p = 1, 2, ..., p, \tag{31}$$

where ρ^2 represents the correlation between the explanatory variables and w_{ij}'s are independent standard normal pseudo-random numbers. The response variable is generated by

$$y_i = \beta_0 + \beta_1 x_{i1} + ... + \beta_p x_{ip} + \varepsilon_i, \tag{32}$$

where ε_i is independent and identically normal distributed pseudo-random numbers with zero mean and variance σ^2. Because the sample size has direct impact on the prediction accuracy, three representative values of the sample size are considered: 30, 50 and 150. In addition, the number of the explanatory variables are considered as $p \in \{4, 8, 12\}$. Further, because we are interested in the effect of multicollinearity, in which the degrees of correlation considered more important, three values of the pairwise correlation are considered with $\rho = \{0.90, 0.95, 0.99\}$. Besides, the values of σ^2 is 1.

For a combination of these different values of n, p, ρ, the generated data is repeated 5000 times and the averaged MSE is calculated as

$$MSE(\hat{\beta}) = \frac{1}{5000} \sum_{i=1}^{5000} (\hat{\beta} - \beta)^T (\hat{\beta} - \beta), \tag{33}$$

where $\hat{\beta}$ is the obtained ridge estimator by the methods. The MSE values from the Monte Carlo simulation study are reported in Tables 1 – 3. Several observations can be obtained as follows:

1. The simulation results indicate that the HHOA method of selecting K is superior to the other used selection methods for all combinations of n, p, and ρ in terms of MSE. We can see that HHOA method has smaller MSE and significantly lower MSE than others.
2. It is seen from Tables 1 -3 that $\hat{\beta}_{HHOA}$ estimator using HHOA method is usually more efficient than the OLS estimator for all values of n, p and when multicollinearity is high or severe.
3. In terms of ρ values, there is increasing in the MSE values when the correlation degree increases regardless the value of n and p.
4. Regarding the number of explanatory variables, it is easily seen that there is a negative impact MSE, where there are increasing in their values when the p increasing from four explanatory variables to twelve explanatory variables.
5. With respect to the value of n, the MSE values decrease when n increases, regardless the value of $\rho$and p.
6. All the selection methods of $K$are superior to the OLS estimator in terms of MSE.

5 Real application results

To evaluate the predictive performance of the proposed method and to compare its performance with the other used methods in a real data application, the Portland cement dataset is employed. Portland cement dataset became a standard dataset to examine and to remedy the multicollinearity. It was widely used by numerous researchers. This

Table 1: Average MSE when $n = 50$

		OLS	HHOA	HK	KN	TC	f	HSL	AH	D	SB	DK	SV1	SV2	AS	M
p=4	r=0.90	1.926	0.912	1.006	1.044	0.976	0.93	0.947	0.947	0.941	0.976	1.006	1.028	1.028	1.007	0.982
	r=0.95	1.934	0.924	0.996	1.022	0.978	0.95	0.964	0.964	0.952	0.978	0.996	1.015	1.015	0.998	0.982
	r=0.99	1.918	0.932	0.959	0.968	0.947	0.937	0.96	0.96	0.935	0.947	0.958	0.971	0.971	0.969	0.96
p=8	r=0.90	2.826	0.732	0.929	1.008	0.874	0.829	0.839	0.839	0.84	0.874	0.928	0.992	0.992	0.931	0.891
	r=0.95	2.835	0.755	0.923	0.97	0.885	0.85	0.869	0.869	0.851	0.885	0.922	0.968	0.968	0.928	0.893
	r=0.99	2.819	0.762	0.905	0.927	0.87	0.846	0.888	0.888	0.845	0.87	0.904	0.945	0.945	0.928	0.902
p=12	r=0.90	3.765	0.711	0.894	0.98	0.829	0.774	0.797	0.797	0.783	0.829	0.894	0.975	0.975	0.897	0.831
	r=0.95	3.767	0.720	0.889	0.959	0.826	0.785	0.815	0.815	0.78	0.826	0.889	0.969	0.969	0.899	0.838
	r=0.99	3.755	0.736	0.86	0.896	0.806	0.772	0.855	0.856	0.773	0.806	0.859	0.94	0.94	0.906	0.859

Table 2: Average MSE when $n = 150$

		OLS	HHOA	HK	KN	TC	f	HSL	AH	D	SB	DK	SV1	SV2	AS	M
p=4	r=0.90	1.942	0.913	1.008	1.034	0.984	0.943	0.946	0.946	0.952	0.984	1.008	1.022	1.022	1.008	0.991
	r=0.95	1.982	0.923	1.013	1.029	1	0.983	0.994	0.994	0.987	1.001	1.013	1.022	1.022	1.013	1.006
	r=0.99	1.967	0.959	0.983	0.988	0.978	0.971	0.983	0.983	0.971	0.978	0.983	0.987	0.987	0.983	0.981
p=8	r=0.90	2.948	0.848	0.998	1.032	0.975	0.948	0.957	0.957	0.953	0.975	0.998	1.021	1.021	0.998	0.978
	r=0.95	2.939	0.871	0.971	1.003	0.955	0.94	0.951	0.951	0.942	0.955	0.971	0.994	0.994	0.971	0.956
	r=0.99	2.946	0.875	0.974	0.989	0.961	0.951	0.975	0.975	0.95	0.961	0.974	0.991	0.991	0.978	0.967
p=12	r=0.90	3.922	0.884	0.975	1.022	0.944	0.922	0.933	0.933	0.924	0.944	0.974	1.012	1.012	0.975	0.944
	r=0.95	3.925	0.917	0.971	1.009	0.946	0.926	0.942	0.942	0.927	0.946	0.971	1.003	1.003	0.972	0.947
	r=0.99	3.913	0.927	0.951	0.974	0.929	0.916	0.947	0.947	0.916	0.929	0.951	0.981	0.988	0.956	0.937

dataset comes from an experimental investigation of heat evolved during the setting and hardening of Portland cements of varied composition and the dependence of this heat on the percentages of four compounds in the clinkers from which the cement was produced. There are 13 observations of heat evolved in calories per gram of cement (y), tricalcium aluminate (x_1), tetracalcium silicate (x_2), tetracalcium alumino ferrite (x_3), and dicalcium silicate (x_4).

It is apparent from Table 4 that there is an improvement of the predictive capability of the HHOA comparing with the other used methods, where HHOA significantly reduced the MSE. The reduction of MSE using HHOA was 11.350%, 10.892%, 10.625%, 10.552%, 10.007%, 10.888%, 9.842%, 11.101%, 9.842%, 11.101%, 9.843%, 10.969%, 9.945%, 10.056%, 10.431%, and 10.213% compared with OLS, HK, KN, TC, f, HSL, AH, D, SB, DK, SV1, SV2, AS , and M, respectively.

Table 3: Real application results for the used methods

Method	MSE
OLS	9303.049
HHOA	8247.127
HK	9255.305
KN	9227.561
TC	9220.055
f	9164.227
HSL	9254.874
AH	9147.469
D	9277.002
SB	9147.553
DK	9263.284
SV1	9157.882
SV2	9169.254
AS	9207.598
M	9185.278

6 Conclusion

In this paper, a new shrinkage parameter selection of the generalized ridge estimator, which is depending on employing the Harris hawks optimization algorithm , was proposed. This proposed method allows us to handle multicollinearity with decreasing the variability of shrinkage parameter selection. Simulation and results demonstrate that the proposed method is outperformed several classical methods. Furthermore, the results proved that the proposed method is more efficient than Hoerl and Kennard (1970).

References

Al-Taweel, Y. and Algamal, Z. (2020). Some almost unbiased ridge regression estimators for the zero-inflated negative binomial regression model. *Periodicals of Engineering and Natural Sciences*, 8(1):248–255.

Algamal, Z. Y. (2018a). Developing a ridge estimator for the gamma regression model. *Journal of Chemometrics*, 32(10):e3054.

Algamal, Z. Y. (2018b). A new method for choosing the biasing parameter in ridge estimator for generalized linear model. *Chemometrics and Intelligent Laboratory Systems*, 183:96–101.

Algamal, Z. Y. (2018c). Shrinkage estimators for gamma regression model. *Electronic Journal of Applied Statistical Analysis*, 11(1):253–268.

Algamal, Z. Y. (2020). Shrinkage parameter selection via modified cross-validation approach for ridge regression model. *Communications in Statistics-Simulation and Computation*, 49(7):1922–1930.

Algamal, Z. Y. and Alanaz, M. M. (2018). Proposed methods in estimating the ridge regression parameter in poisson regression model. *Electronic Journal of Applied Statistical Analysis*, 11(2):506–515.

Algamal, Z. Y. and Asar, Y. (2020). Liu-type estimator for the gamma regression model. *Communications in Statistics-Simulation and Computation*, 49(8):2035–2048.

Alheety, M. and Kibria, B. G. (2014). A generalized stochastic restricted ridge regression estimator. *Communications in Statistics-Theory and Methods*, 43(20):4415–4427.

Alkhamisi, M. A. and Shukur, G. (2007). A monte carlo study of recent ridge parameters. *Communications in Statistics—Simulation and Computation®*, 36(3):535–547.

Alkhateeb, A. and Algamal, Z. (2020). Jackknifed liu-type estimator in poisson regression model. *Journal of The Iranian Statistical Society*, 19(1):21–37.

Alobaidi, N. N., Shamany, R. E., and Algamal, Z. Y. (2021). A new ridge estimator for the negative binomial regression model. *Thailand Statistician*, 19(1):116–125.

Asar, Y., Karaibrahimoğlu, A., and Genç, A. (2014). Modified ridge regression parameters: A comparative monte carlo study. *Hacettepe Journal of Mathematics and Statistics*, 43(5):827–841.

Batah, F. S. M., Ramanathan, T. V., and Gore, S. D. (2008). The efficiency of modified jackknife and ridge type regression estimators: a comparison. *Surveys in Mathematics & its Applications*, 3.

Bhat, S. and Raju, V. (2017). A class of generalized ridge estimators. *Communications in Statistics-Simulation and Computation*, 46(7):5105–5112.

Bhat, S. S. (2016). A comparative study on the performance of new ridge estimators. *Pakistan Journal of Statistics and Operation Research*, 12(2):317–325.

Dorugade, A. (2014). New ridge parameters for ridge regression. *Journal of the Association of Arab Universities for Basic and Applied Sciences*, 15(1):94–99.

Dorugade, A. and Kashid, D. (2010). Alternative method for choosing ridge parameter for regression. *Applied Mathematical Sciences*, 4(9):447–456.

Firinguetti, L. (1999). A generalized ridge regression estimator and its finite sample properties: A generalized ridge regression estimator. *Communications in Statistics-Theory and Methods*, 28(5):1217–1229.

Hameed, S. S., Hassan, R., and Muhammad, F. F. (2017). Selection and classification of gene expression in autism disorder: Use of a combination of statistical filters and a gbpso-svm algorithm. *PLoS One*, 12(11):e0187371.

Heidari, A. A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M., and Chen, H. (2019). Harris hawks optimization: Algorithm and applications. *Future generation computer systems*, 97:849–872.

Hocking, R. R., Speed, F., and Lynn, M. (1976). A class of biased estimators in linear regression. *Technometrics*, 18(4):425–437.

Hoerl, A. E. and Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67.

Nomura, M. (1988). On the almost unbiased ridge regression estimator. *Communications in Statistics-Simulation and Computation*, 17(3):729–743.

Rashad, N. K. and Algamal, Z. Y. (2019). A new ridge estimator for the poisson regression model. *Iranian Journal of Science and Technology, Transactions A: Science*, 43(6):2921–2928.

Shamany, R., Alobaidi, N. N., and Algamal, Z. Y. (2019). A new two-parameter estimator for the inverse gaussian regression model with application in chemometrics. *Electronic Journal of Applied Statistical Analysis*, 12(2):453–464.

Troskie, C. and Chalton, D. Detection of outliers in the presence of multicollinearity. In *Multidimensional statistical analysis and theory of random matrices, Proceedings of the Sixth Lukacs Symposium, eds. Gupta, AK and VL Girko*, pages 273–292.

Yahya Algamal, Z. (2019). Performance of ridge estimator in inverse gaussian regression model. *Communications in Statistics-Theory and Methods*, 48(15):3836–3849.

Yang, S.-P. and Emura, T. (2017). A bayesian approach with generalized ridge estimation for high-dimensional regression and testing. *Communications in Statistics-Simulation and Computation*, 46(8):6083–6105.