AUT Journal of Mathematics and Computing

AUT J. Math. Comput., 6(4) (2025) 353-360
https://doi.org/10.22060/AJMC.2024.22963.1208

Original Article

SCAD regression model selection with information criteria for multivariate response models

Amir Hossein Ghatari[a], Mina Aminghafari[a], Omid Naghshineh Arjmand[*a]

[a] *Department of Statistics, Faculty of Mathematics and Computer Science, Amirkabir University of Technology, Tehran, Iran*

ABSTRACT: This paper provides an objective function for smoothly clipped absolute deviation (SCAD) regression models with multivariate responses. The log-likelihood of a multivariate normal distribution is considered instead of L_2 norm to create the model's objective function. Additionally, the SCAD penalty has a tuning parameter, and the information criteria, suitable for the proposed model are presented to select the tuning parameter. Based on numerical studies, the consistency of the proposed information criteria is checked via simulation experiments. Moreover, the best criterion is introduced using simulated and real datasets.

Review History:

Received:03 February 2024
Revised:30 June 2024
Accepted:01 July 2024
Available Online:01 October 2025

Keywords:

Consistency
Multivariate response
SCAD penalty
Tuning parameter

MSC (2020):

62J05; 62J12

1. Introduction

Penalized regression stands as a potent technique for modeling high-dimensional data, harmonizing traditional regression with penalty terms that foster sparsity or endorse other attributes in the estimated coefficients. This amalgamation facilitates efficient variable selection and simultaneous parameter estimation, even when multicollinearity is prevalent.

SCAD regression, introduced by [4], has numerous advantages. It stimulates sparsity in the estimated coefficients via a penalty mechanism that pushes many coefficients to zero. This feature enables automatic variable selection by discarding insignificant predictors, thereby yielding more concise models. Moreover, SCAD regression adeptly manages multicollinearity, similar to other penalized techniques. Notably, SCAD regression demonstrates asymptotic oracle properties, ensuring consistent estimation of the true model, even in instances where the number of predictors surpasses the sample size.

The literature encompassing SCAD regression is rich in significant references. Fan and Li [4] inaugurated the SCAD penalty, underscoring its non-concave penalized likelihood and oracle characteristics. Fan and Peng [5]

*Corresponding author.
E-mail addresses: a.h.ghatari@aut.ac.ir, naghshineh@aut.ac.ir, aminghafari@aut.ac.ir

extended SCAD regression to accommodate an extensive array of predictors, establishing the variable selection consistency under specific conditions. Zhang et al. [17] delved into the theoretical aspects of SCAD regression, highlighting the consistency of unbiased variable selection and estimation. Additionally, [9] and [10] concentrated on variable selection in high-dimensional data utilizing penalty methods, presenting theoretical findings and practical guidelines for model estimation.

Regression models featuring multivariate responses (multi-responses) confer a multidimensional facet to statistical modeling. Managing multi-response scenarios presents the core challenge. One approach initially disregards the multi-response nature, treating responses as separate univariate regressions. However, this approach eliminates the correlations among multi-response elements. Breiman and Friedman [2] outlined a strategy to address multi-response predicaments via individual regressions. For penalized models, [3] explored diverse variable selection methods for multi-response challenges in their Ph.D. thesis. Variyath and Brobbey [15] further extended this investigation and included selection criteria along with examples involving more than two dimensions for the problem.

Recently, [11] worked on statistical inference about multi-response regression models via providing confidence intervals. [13] applied joint models for multi-response longitudinal data in Bayesian literature. Additionally, [8] proposed multivariate form of Bridge regression and demonstrated the better performance of their selection criterion for high-dimensional data. While the extensive examination of SCAD regression pertains primarily to univariate response data, its application to multivariate response regression remains a subject of ongoing research.

This paper introduces, for the first time, a specialized objective function for SCAD regression tailored to multi-response problems. We also propose four novel information criteria based on generalized information criterion definitions to facilitate the selection of tuning parameters. Through a simulation study, we assess the consistency of the models chosen by these four criteria. Additionally, we evaluate the performance of SCAD regression for multi-response scenarios using both simulated and real datasets. Section 2 reviews the essential prerequisites. Section 3 presents the objective function of SCAD regression for multi-response problems and the new information criteria. Finally, Section 4 provides numerical analyses that demonstrate our findings and results.

2. Basic Concepts

This section reviews the basic definitions required to obtain the paper's main results.

2.1. Multi-response Penalized Regression

The multi-responses regression model is defined as follows:

$$Y_{n\times K} = X_{n\times p} B_{p\times K} + E_{n\times K}, \tag{1}$$

where $Y = [Y_1 \,|Y_2| \cdots |\, Y_K]$ is the observation matrix of the multi-response with K elements, X is the feature observation and $B = [\beta_1 \,|\beta_2| \cdots |\, \beta_K]$ is the matrix of regression coefficients. $E = [E_1 \,|E_2| \cdots |\, E_K]$ is the matrix of noise. For $j = 1, \ldots, K$, we have $Y_j = (y_{1j}, \ldots, y_{nj})'$, $\beta_j = (\beta_{1j}, \ldots, \beta_{pj})$ and $E_j = (\varepsilon_{1j}, \ldots, \varepsilon_{nj})'$ are the jth columns of Y, B and E, respectively.

The general form of the objective function in penalized models is considered as:

$$Q(\lambda, B) = \ell(B) + p\,(\lambda, |B|)\,, \tag{2}$$

where $\ell(.)$ and $p\,(.)$ are the negative log-likelihood and penalty functions, respectively. Additionally, λ is the tuning parameter of the penalized model. There are several penalty functions, such as bridge, proposed by [6], LASSO, proposed by [14], and SCAD penalty function, proposed by [4]. After identifying the objective function, providing a method to select the optimum λ is the next step. Cross-validation and information criteria are tools for obtaining λ via this approach.

2.2. Information Criteria

Information criteria can be used to select penalized or non-penalized models. For example, the well-known AIC has both non-penalized [1] and penalized [18] forms for model selection. All information criteria mostly have two parts: a function of likelihood and a penalty part. For models with multi-response, [15] proposed their criteria, GCV and BIC, as:

$$GCV(\lambda) = \frac{1}{n}\frac{D}{\left(1 - n^{-1} df_\lambda\right)^2}, \tag{3}$$

and

$$BIC(\lambda) = \log\left(\frac{D}{n}\right) + \left(\frac{\log(n)}{n}\right) df_\lambda, \tag{4}$$

where D is the deviance of the model obtained from the null model (a model without features), and df_λ is the number of features in the model. We use the GCV and BIC in Section 4 for comparison study. We propose a type of information criterion to select λ, based on the studies of [18] and [7], in Section 3.

3. Multi-response SCAD Regression

In this section, we propose the objective function of SCAD regression, which is suitable for multi-response models, and then we specify the information criteria, proposed by [16], [18], and [7] for the provided SCAD regression.

3.1. Objective Function of Multi-SCAD

It is required to clarify objective function (2) in this situation. The objective function has two parts. $\ell(B)$ and a penalty function. We should determine these parts for a problem with multivariate normal noise and the SCAD penalty function. For multi-response models with normal noise, we have

$$\ell(B) = \frac{1}{Kn}\sum_{h=1}^{n}\left(Y^{(h)} - B_\lambda^T X_h\right)^T \Sigma^{-1}\left(Y^{(h)} - B_\lambda^T X_h\right),$$

where $Y^{(h)}$ and X_h are the hth row and columns of Y and X respectively. On the other hand, [4] proposed the following SCAD penalty function:

$$Pe_\lambda(\beta;\alpha) = \begin{cases} \lambda|\beta| & |\beta| \le \lambda \\ -\left(\beta^2 - 2\alpha\lambda|\beta| + \lambda^2\right)/[2(a-1)] & \lambda \le |\beta| \le \alpha\lambda \\ (a+1)\lambda^2/2 & |\beta| \ge \alpha\lambda \end{cases}$$

Also, they suggested $\alpha = 3.7$ as an optimum value for α. We consider the SCAD penalty for choosing any element of B (β_{ij}, $i = 1, \dots, p$ and $j = 1, \dots, K$) as:

$$Pe_\lambda(\beta_{ij}) = \begin{cases} \lambda|\beta_{ij}| & |\beta_{ij}| \le \lambda \\ -\left(\beta_{ij}^2 - 7.4\lambda|\beta_{ij}| + \lambda^2\right)/5.4 & \lambda \le |\beta_{ij}| \le 3.7\lambda \\ 2.35\lambda^2 & |\beta_{ij}| \ge 3.7\lambda \end{cases} \tag{5}$$

Finally, we can consider the objective function for multi-response SCAD regression as follows:

$$Q(B_\lambda, \lambda) = \frac{1}{Kn}\sum_{h=1}^{n}\left(Y^{(h)} - B_\lambda^T X_h\right)^T \Sigma^{-1}\left(Y^{(h)} - B_\lambda^T X_h\right) + \sum_{i=1}^{P}\sum_{j=1}^{K} Pe_\lambda\left(\beta_{ij}\right) \tag{6}$$

Note that based on the structure of our penalty function (5), we specialized the SCAD penalty for every single element of B matrix (β_{ij},$i = 1, \dots, p$ & $j = 1, \dots, K$) and in the objective function (6), we have a double summation on (5) to obtain the penalty of B. Indeed, we can consider a SCAD structure for each β_{ij}. Hence, it is reasonable to follow what the providers of univariate SCAD did for α.

The penalized estimation of B is called $\hat{B}_\lambda$ and is obtained by minimizing objective function (6). The next step is to present the methods for selecting λ. In this paper, we use cross-validation and information criteria to obtain λ. We propose the following information criterion which is applicable to multi-response SCAD regression.

3.2. GIC Types for Tunning Parameter Selection

Choosing the optimum tuning parameter λ is one of the main goals of penalized models. Based on the general format of the generalized information criterion GIC, pioneered by [18], we introduce the multivariate generalized information criterion (MGIC) as:

$$\text{MGIC}(\lambda) = G(Y, \hat{B}_\lambda) + \frac{k_n df_\lambda}{Kn}, \tag{7}$$

where Y is the observation matrix of multi-response, $\hat{B}_\lambda$ is the estimator of B for a fixed λ and $G(Y, \hat{B}_\lambda)$ measures the fitness of the model. [18] divided the GIC family into two types: AIC-type and BIC-type. When $k_n \longrightarrow \infty$ ($n \longrightarrow \infty$) and $\frac{k_n}{\sqrt{n}} \longrightarrow 0$, we have the BIC-type, and if $k_n \longrightarrow 2$, we have AIC-type criteria.

If we have a residual matrix $e_{\lambda_{n\times K}} = Y - X\hat{B}_\lambda$, then we can consider $\log(\det(S_{e_\lambda}))$ instead of G as a tool that measures the fitness of the model. S_{e_λ} is the estimation for the covariance matrix of e_λ.

Now, we consider $k_n = \log(n)$, which leads to multivariate BIC-type of (7) (called MBIC) as:

$$\text{MBIC}_1(\lambda) = \log(\det(S_{e_\lambda})) + \frac{\log(n) df_\lambda}{Kn}.$$

If we set $k_n = \sqrt[3]{n}$, we can define another type of MBIC as:

$$\text{MBIC}_2(\lambda) = \log(\det(S_{e_\lambda})) + \frac{\sqrt[3]{n} df_\lambda}{Kn},$$

which satisfies the conditions of the BIC-type. If we define

$$G^*(Y, \hat{B}_\lambda) = \frac{1}{n}\sum_{i=1}^{n} e_{\lambda_i}^T S_{e_\lambda}^{-1} e_{\lambda_i},$$

where e_i is the ith row of e_λ. This type of G considers the correlation among the columns of e_λ. G^* is defined based on the Mahalanobis norm of e_λ columns. We use this definition to include the correlation among the columns of e_λ. Additionally, $\det(S_{e_\lambda})$ considers the correlation in another way. Indeed, we define G^* to have another interpretation of the correlation in the structure of the G function.

Based on the proposed form of AIC-type by [18] and G^*, we define multivariate AIC-type of (7) (MAIC) as follows:

$$\text{MAIC}(\lambda) = \log(G^*(Y, \hat{B}_\lambda)) + \frac{2G^*(Y, \hat{B}_\lambda) df_\lambda}{Kn}.$$

Also, [7] proposed a new form of AIC-type for GIC called AGIC. Using the form of $AGIC$, we remove G^* in the penalty part of the criteria and define the multivariate form of AGIC (MAGIC) as:

$$\text{MAGIC}(\lambda) = \log\left(\frac{1}{K}\sum_{i=1}^{K} \hat{\sigma}_i^2\right) + \frac{2df_\lambda}{Kn}.$$

The λ value that minimizes each criterion is called the tuning parameter selected by the criterion. Moreover, $\hat{B}_\lambda$ obtained by the selected λ is the penalized estimation of B presented by the criterion.

4. Numerical Experiments

We study the performance of the proposed criteria, for the MGIC family including MAIC, MAGIC, MBIC_1, and MBIC_2, via simulation experiments and analysis of real dataset. First, we check the consistency of the selected model based on the proposed criteria. Then, we compare the performance of the MGIC family with BIC, GCV, (see (3) and (4)) and CV method.

4.1. Simulation Study

We consider model (1) with objective function (6) in the simulation study. We generate data for bivariate response cases. We have the following structure to generate data:

$$\begin{aligned} &Y = XB + E, \\ &X \sim N_p(0, \Gamma), \quad \Gamma = \left[(\rho_X)^{i-j}\right]_{p\times p}, \; E \sim N_2(0, \Sigma), \end{aligned} \tag{8}$$

where $\Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$ and $B = [B_1 \mid B_2]$.

Additionally, $B_1 = \left(3, 3, 3, 3, 3, \vec{O}_{(p-5)\times 1}\right)^T$, $B_2 = \left(4, 4, 4, 4, 4, 4, 4, 4, 4, 4, \vec{O}_{(p-10)\times 1}\right)^T$, and $\vec{O}$ is a vector with zero elements. If at least one of the feature parameters becomes nonzero, it is considered a relevant feature. Note the number of relevant features is 10.

4.1.1. Consistency

To show the consistency of the estimator, we verify whether the MSE tends to zero. We compute a sequence of MSEs as follows:

$$MSE(\hat{B}_{\lambda_i}) = \frac{1}{i}\sum_{k=1}^{K}\sum_{j=l}^{p}(\beta_{kj} - \hat{\beta}^i_{kj})^2, \quad i = 1, \dots, n,$$

where λ_i is selected λ and $\hat{\beta}^i_{kj}$ is the estimate of β_{kj} obtained from a sample with a size of i. We want to check whether $\lim_{n\to\infty} MSE\left(\hat{B}_{\lambda_n}\right) = 0$ or not. For this purpose, we consider the following definition of the limit of a sequence:

$$\forall \varepsilon > 0 \quad \exists \ell \in \mathbb{N} \quad \text{such that} \quad \forall k \geq \ell, \quad \left|MSE(\hat{B}_{\lambda_k})\right| < \varepsilon,$$

where ℓ is the convergence point. We use this structure to verify the consistency of the MGIC family. We consider $\rho_X = 0.25, 0.5, 0.75$ and $p = 40$, and we generate the simulated data with a size of $n = 10000$. Note that if $\ell = n$, we conclude that the sequence has not converged over n observations.

For a bivariate form of multi-SCAD regression, we calculate the MSE sequence of $\hat{B}_{\lambda_n}$ obtained by four criteria. We use the average of the MSE sequences over 25 iterations.

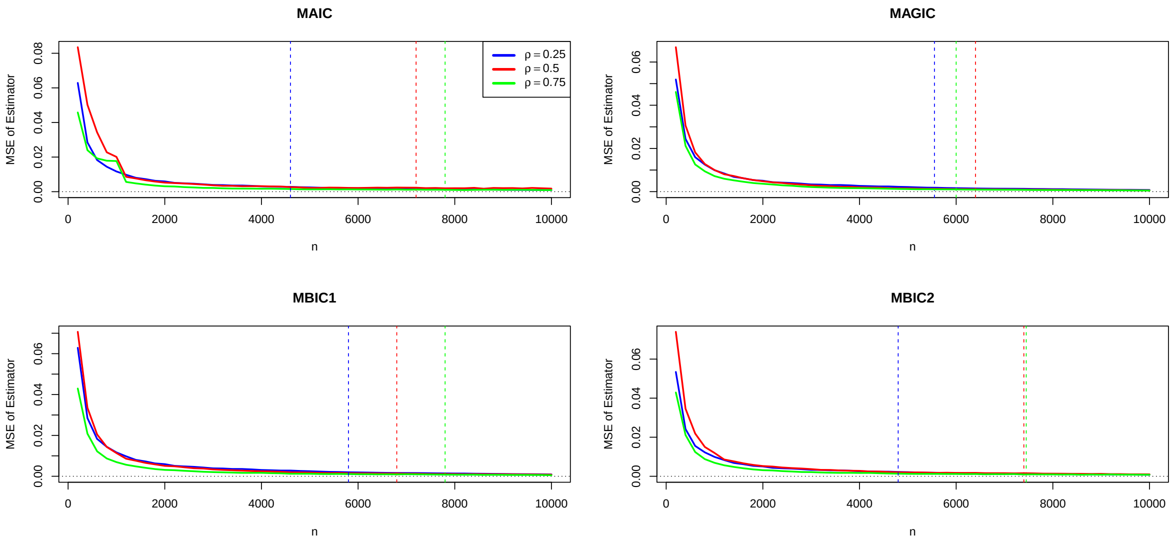

Figure 1: The average sequences of MSE for different criteria and values of ρ. The vertical dashed lines show the convergence points of each curve.

Figure 1 shows the average curves for the MSE sequences for different values of ρ_X and the proposed criteria. The vertical dashed lines show the convergence points of each curve. All the sequences converged before n. Moreover, the sequence converged faster when $\rho_X = 0.25$ for all criteria. The reason for this faster converging is that feature matrix X has no collinearity. As is well-known, penalization has several aims, such as removing collinearity and selecting features. In the absence of collinearity, it is reasonable that a penalized model performs better.

4.1.2. Comparison Study

The performance of the MGIC family was studied from the viewpoint of consistency. We want to compare their performance to GCV, BIC, and CV. We consider the bivariate form of model (8) and $\rho_X = 0.25, 0.5$ and 0.75 and generate data with sample sizes $n = 200, 400$, $p = 40, 120$, and 1000 iterations. We consider a test dataset with the size of $n_{test} = 100$ for each situation.

Figure 2 includes box plots of the number of the selected features for different values of ρ_X. The increase in n leads to the selection of fewer features. Increasing n leads to the removal of irrelevant features by the MGIC family due to their consistency. This approach is compatible with Parsimony's principle. Also, the MAIC criteria select the minimum number of features for each situation. For all MGIC variants, the main boxes are strictly around 10. Additionally, there is a significant decrease in the number of selected features where ρ_X increases. CV selects the greatest number of features.

In what follows, we analyze the performance of the studied criteria based on the error in the test datasets. Instead of the usual form of MSE in the regression model, we consider MSE_{MH} using the Mahalanobis distance, which is more suitable for multi-SCAD objective function (6). The MSE_{MH} is calculated as:

$$MSE_{MH} = \frac{1}{n}\sum_{i=1}^{n} e_i' S_e^{-1} e_i,$$

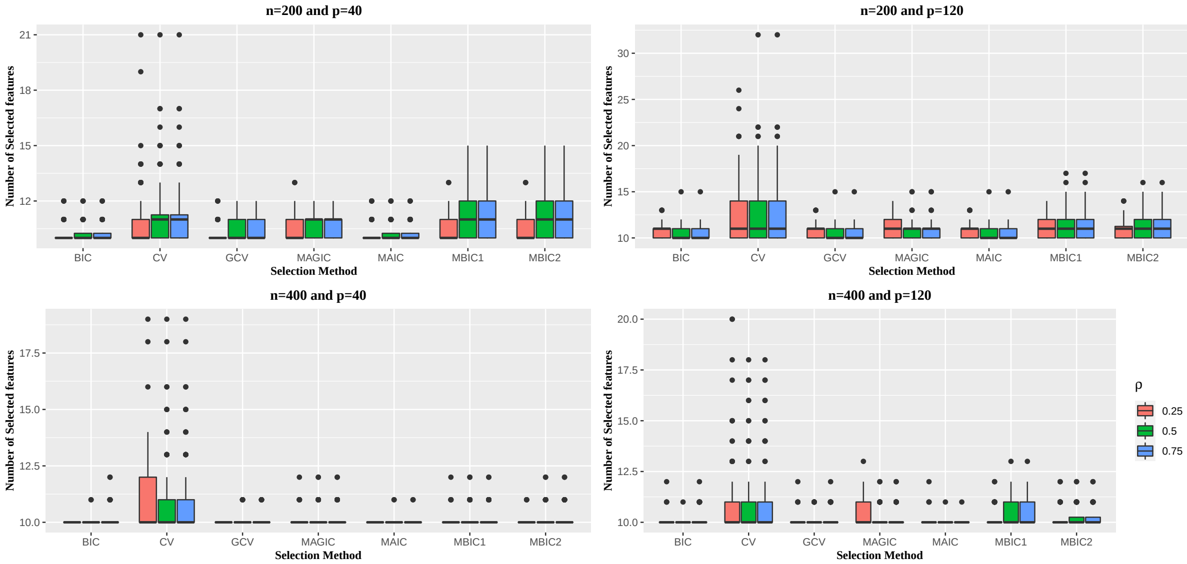

Figure 2: Box plot of the selected features for $p = 40, 120$ and $n = 200, 400$ and different values of ρ_X.

where e_i is the ith row of e_λ (the residual matrix of the model), and S_{e_λ} is the sample covariance matrix of e_λ. Note that the lower MSE_{MH} concludes better performance.

Table 1: Means of $MSE_{MH_{test}}$ for selected models by various training datasets over 1000 iterations. Bold values demonstrate the minimum of $MSE_{MH_{test}}$ for each situation.

ρ_X	n	p	MAGIC	MAIC	MBIC$_1$	MBIC$_2$	CV	GCV	BIC
0.25	200	40	1.11	**1.09**	1.15	1.14	1.42	1.22	1.17
		120	1.15	**1.11**	1.17	1.17	1.45	1.33	1.23
	400	40	**1.12**	**1.12**	1.20	1.21	1.50	1.25	1.27
		120	1.15	**1.12**	1.19	1.22	1.47	1.42	1.24
0.50	200	40	1.20	**1.11**	1.22	1.25	1.40	1.28	**1.11**
		120	1.17	**1.12**	1.19	1.23	1.35	1.19	1.14
	400	40	1.07	**1.05**	1.14	1.14	1.35	1.10	1.09
		120	1.14	**1.10**	1.17	1.18	1.41	1.24	**1.10**
0.75	200	40	1.22	**1.19**	1.25	1.31	1.44	1.27	**1.19**
		120	1.24	**1.20**	1.27	1.31	1.45	1.37	1.24
	400	40	**1.10**	1.21	1.23	1.29	1.31	1.31	**1.10**
		120	**1.12**	**1.12**	1.25	1.22	1.33	1.15	**1.12**

Table 1 shows the average values of the $MSE_{MH_{test}}$ for all possible situations over 1000 iterations. In most cases, the MAIC has the minimum MSE_{MH}. However, increasing n improves the performance of the methods, and increasing ρ has no meaningful effect on the behavior of the methods. MAIC generally performs best from both viewpoints (the number of selected features and MSE_{MH}). This approach involves removing irrelevant features and keeping relevant ones. On the other hand, it has the minimum MSE_{MH} compared to the other methods.

4.2. Real Data Analysis

We study the performance of the MGIC family and others as variable selection methods on the residential building dataset[1]. The dataset includes 102 features and a bivariate response. The construction costs and sale prices make the bivariate response corresponding to single-family residential apartments in Tehran, Iran. The details about the features can be found in [12].

Figure 3 shows the relationship between construction costs (Y_1) and sale prices (Y_2). As it is shown in Figure 3 and according to the results of Pearson's correlation test, the hypothesis of correlation between Y_1 and Y_2 is not rejected.

[1] `https://archive.ics.uci.edu/ml/datasets/Residential+Building+Data+Set`

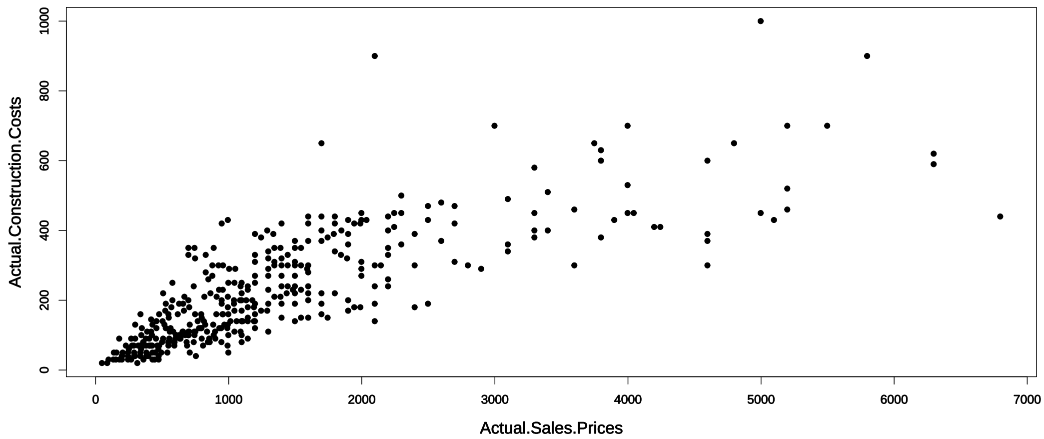

Figure 3: The values of construction cost versus sale prices.

Table 2: Summarized model selection results for the residential building dataset. Bold values show the minimum $MSE_{MH_{test}}$.

Values	MAGIC	MAIC	MBIC_1	MBIC_2	CV	GCV	BIC
p^*	11	12	11	12	15	15	14
$MSE_{MH_{test}}$	22.11	**16.21**	22.11	22.11	23.34	22.89	19.83

Table 2 shows the results of applying the MGIC family and other criteria. p^* indicates the number of selected features. Like the simulation study, MAIC has the minimum MSE_{MH}, whereas it selects 12 features more than MAGIC and MBIC_1. However, they have greater MSE_{MH} than MAIC.

Conclusion

Multivariate responses present challenges in regression problems, as the potential correlations among the elements of the response variables cannot be disregarded. This paper proposes utilizing the Mahalanobis norm instead of the conventional L_2 norm in the objective function to address these correlations. Another challenge lies in selecting the tuning parameter within penalized models. To this end, we introduce a specialized generalized information criterion (GIC) format tailored for selecting these parameters. Although a closed-form solution for estimating the coefficient vector in SCAD regression remains elusive, our study verifies the consistency of the estimated coefficients generated by these criteria through comprehensive simulation studies. Additionally, the performance of SCAD regression with multi-response problems is demonstrated using both simulated and real datasets. The results show that the Mahalanobis norm effectively accounts for correlations among response variables, and the proposed GIC format reliably selects tuning parameters for penalized models. These findings support the potential of our approach to improve the accuracy and interpretability of multivariate regression models. For future research, it would be valuable to provide a structure to select consistent models in the concept of non-concave penalty functions. Additionally, providing a new version of multi-SCAD for models with non-Gaussian noise can be considered as a new subject.

References

[1] H. Akaike, *A new look at the statistical model identification*, IEEE Transactions on Automatic Control, 19 (1974), pp. 716–723.

[2] L. Breiman and J. H. Friedman, *Predicting multivariate responses in multiple linear regression*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 59 (1997), pp. 3–54.

[3] A. Brobbey, *Variable selection in multivariate multiple regression*, PhD thesis, Memorial University of Newfoundland, 2015.

[4] J. Fan and R. Li, *Variable selection via nonconcave penalized likelihood and its oracle properties*, Journal of the American Statistical Association, 96 (2001), pp. 1348–1360.

[5] J. Fan and H. Peng, *Nonconcave penalized likelihood with a diverging number of parameters*, The Annals of Statistics, 32 (2004), pp. 928–961.

[6] L. E. Frank and J. H. Friedman, *A statistical view of some chemometrics regression tools*, Technometrics, 35 (1993), pp. 109–135.

[7] A. H. Ghatari and M. Aminghafari, *A new type of generalized information criterion for regularization parameter selection in penalized regression*, Journal of Biopharmaceutical Statistics, (2023), pp. 1–25.

[8] ——, *Multi-response bridge regularization parameter selection via multivariate generalized information criterion*, Fluctuation and Noise Letters, (2024).

[9] X. Guo, H. Zhang, Y. Wang, and J.-L. Wu, *Model selection and estimation in high dimensional regression models with group scad*, Statistics & Probability Letters, 103 (2015), pp. 86–92.

[10] T. Honda and C.-T. Lin, *Forward variable selection for ultra-high dimensional quantile regression models*, Annals of the Institute of Statistical Mathematics, 75 (2023), pp. 393–424.

[11] Y. Ma, Y. Li, and J. Xu, *Confidence intervals for high-dimensional multi-task regression*, JUSTC, 53 (2023), pp. 1–9.

[12] M. H. Rafiei and H. Adeli, *A novel machine learning model for estimation of sale prices of real estate units*, Journal of Construction Engineering and Management, 142 (2016), pp. 1–10.

[13] Y.-Z. Tian, M.-L. Tang, C. Wong, and M.-Z. Tian, *Bayesian analysis of joint quantile regression for multi-response longitudinal data with application to primary biliary cirrhosis sequential cohort study*, Statistical Methods in Medical Research, (2024), pp. 1–22.

[14] R. Tibshirani, *Regression shrinkage and selection via the lasso*, Journal of the Royal Statistical Society. Series B (Methodological), (1996), pp. 267–288.

[15] A. M. Variyath and A. Brobbey, *Variable selection in multivariate multiple regression*, Plos One, 15 (2020), p. e0236067.

[16] H. Wang and C. Leng, *Unified lasso estimation by least squares approximation*, Journal of the American Statistical Association, 102 (2007), pp. 1039–1048.

[17] C.-H. Zhang, *Nearly unbiased variable selection under minimax concave penalty*, The Annals of Statistics, 38 (2010), pp. 894–942.

[18] Y. Zhang, R. Li, and C.-L. Tsai, *Regularization parameter selections via generalized information criterion*, Journal of the American Statistical Association, 105 (2010), pp. 312–323.

Please cite this article using:

Amir Hossein Ghatari, Mina Aminghafari, Omid Naghshineh Arjmand, SCAD regression model selection with information criteria for multivariate response models, AUT J. Math. Comput., 6(4) (2025) 353-360
https://doi.org/10.22060/AJMC.2024.22963.1208