Laplace LIC for Distributed Estimation

Yaxuan Wang, Guangbao Guo, Weidong Wu

***Abstract*—The Laplace regression model is widely used in statistical analysis, assuming that the error terms follow a Laplace distribution. To address the issue of information redundancy, we utilize the proposed Laplace LIC criterion to select the optimal subset from the large-scale data. The Laplace LIC criterion is derived based on minimizing the confidence interval length and maximizing the determinant of information matrix. Experimental analysis demonstrates that the Laplace LIC criterion has robust stability.**

***Index Terms*—Laplace LIC criterion, distributed estimation, optimal subset.**

I. Introduction

IN the field of statistical analysis, traditional ordinary least squares (OLS) regression assumes that the error term follows a normal distribution. These models are highly sensitive to outliers and lack adaptability to non-normal data. To address this issue, this paper introduces the Laplace regression model. By leveraging the Laplace distribution, the model reduces the influence of outliers and enhances computational efficiency (see [5], [6]). This approach strengthens the models robustness and flexibility, enabling better handling of outliers and non-normal data.

A. Current research status

In statistics, research on redundant data has always been a hot topic (see [2], [11], [12]). This paper proposes a new method based on the Laplace regression model. The model aims to solve the problem of optimal subset selection in redundant data (see [1], [4]). It fills key research gaps, improves computational efficiency and accuracy, and demonstrates good adaptability to outliers and non-normal distributions.

B. Our work

This paper explores the theoretical properties of the Laplace LIC criterion. By applying the LIC criterion, we select the optimal subset from the datasets, where the error terms follow the Laplace distribution. Additionally, we employ the Mean Absolute Error (MAE) and Mean Squared Error (MSE) as performance metrics, and evaluate the LIC criterion's effectiveness through simulation experiments under five common distributions (see [3], [7]).

Finally, we develop an R package named LLIC, built upon the Laplace regression model. The package includes functions for data preprocessing, variable selection, model fitting, and result visualization, enhancing prediction accuracy and improving information utilization in small datasets.

Manuscript received April 10, 2024; revised September 25, 2025.

This work was supported by a grant from National Social Science Foundation Project under project ID 23BTJ059, a grant from Natural Science Foundation of Shandong under project ID ZR2020MA022, a grant from Natural Science Foundation of Shandong under project ID ZR2020QF040, and a grant from National Statistical Research Program under project ID 2022LY016. The authors are with equal contributions.

Yaxuan Wang is an undergraduate student of School of Mathematics and Statistics, Shandong University of Technology, Zibo, China (e-mail: yaxuan041005@163.com).

Guangbao Guo is a professor of School of Mathematics and Statistics, Shandong University of Technology, Zibo, China (corresponding author to provide phone:15269366362; e-mail: ggb11111111@163.com).

Weidong Wu is a lecturer of School of Mathematics and Statistics, Shandong University of Technology, Zibo, China (e-mail: wuweidong.happy@163.com).

II. Interval estimation of distributed Laplace model

In this section, we are interested in the distributed Laplace regression model. It is represented as follows:

$$Y_{I_k} = X_{I_k}\beta + \varepsilon_{I_k}, \varepsilon_{I_k} \sim \text{Laplace}(\mu_1, \text{b}I_{\text{n}_{\text{I}_\text{k}}\times\text{n}_{\text{I}_\text{k}}}),$$

$$k = 1, ..., K_n.$$

X_{I_k} is a $n_{I_k} \times p$ submatrix with $n_{I_k} \geq p$. ε_{I_k} is an error subvector. $I_{n_{I_k}\times n_{I_k}}$ is an unit matrix in $n_{I_k} \times n_{I_k}$. $\beta = (\beta_1, \cdots, \beta_p)^{\text{T}}$ is the regression coefficient vector. μ_1 is a location parameter. b is a scale parameter, which is positive.

For ease of computation, the dataset can be expressed in matrix form:

$$Y = (Y_{I_1}^{\text{T}}, Y_{I_2}^{\text{T}}, \ldots, Y_{I_{K_n}}^{\text{T}})^{\text{T}}, X = (X_{I_1}^{\text{T}}, X_{I_2}^{\text{T}}, \ldots, X_{I_{K_n}}^{\text{T}})^{\text{T}}.$$

Furthermore, the model can be simply represented by the following formula:

$$Y = X\beta + \varepsilon, \varepsilon \sim \text{Laplace}(\mu_1, \text{b}I).$$

The fundamental idea behind distributed estimation is as follows: the massive data on a single computer is distributed across multiple computers. Each computer generates its own local estimator using a statistical inference method [8]. Then, these local estimators are aggregated and averaged to obtain the overall estimator. If the number of blocks for the data is set too large, it may lead to anomalies in some local estimators, which impact the overall estimation result. Here, all the data on a machine is randomly and equally divided into K_n blocks, which are subsequently transmitted to the corresponding computers for processing (see [9], [10]). The K_n subsets are denoted as $Q_{I_k} = (Y_{I_k}, X_{I_k})_{k=1}^{K_n}$. The process of the distributed estimation algorithm is as follows:

1) For $Y_{I_k} = (Y_{I_k,1}, \ldots, Y_{I_k,n_{I_k}})^{\text{T}}$,$k = 1, \ldots, K_n$, the local estimator of the the mean $\mu = E(Y|x)$ is calculated as follows:

$$\hat{\mu}_{I_k} = \overline{Y}_{I_k} = \sum_{i=1}^{n_{I_k}} Y_{I_k,i}/n_{I_k}.$$

2) Aggregate above local estimators and average them to obtain the overall estimator,

$$\hat{\mu}^{(a)} = \frac{1}{K_n}\sum_{i=1}^{K_n}\hat{\mu}_{I_k}.$$

3) The variance for $\hat{\mu}^{(a)}$ is denoted as

$$var(\hat{\mu}^{(a)}) = \frac{1}{K_n^2}\sum_{i=1}^{K_n} var(\hat{\mu}_{I_k}).$$

In order to assess the stability and reliability of the local estimator for each block, we construct a confidence interval $C(Y_{I_k})$ for μ based on Q_{I_k}(shortly, μ_{I_k}), with a specified confidence level of $1-\alpha$. The parameter $w \in (0,1)$ corresponds to a confidence domain function, meaning that each value of w is associated with a specific confidence level. Based on the parameter w, $A_w(\mu_{I_k})$ is denoted as the acceptance region for each μ_{I_k}.

$$A_w(\mu_{I_k}) = \{\overline{Y}_{I_k}\colon (\mu_{I_k} - t_{n_k-p,1-aw}\widehat{\sigma}_{I_k}\cdot\bar{C}_{I_k},$$
$$\mu_{I_k} - t_{n_k-p,\alpha(1-w)}\widehat{\sigma}_{I_k}\cdot\bar{C}_{I_k})\},$$

where $\overline{C}_{I_k} = \sum_{i=1}^{n_{I_k}} C_{I_k,x_i}/n_{I_k}, x_i \in X_{I_k}$, and $C_{I_k,x_i} = x_i^\top (X_{I_k}^\top X_{I_k})^{-1}x_i$ is the diagonal element of the matrix $X_{I_k}(X_{I_k}^\top X_{I_k})^{-1}X_{I_k}^\top$.

Specifically, the confidence interval is derived by inverting the acceptance region at a confidence level α. For example, when $w = 0.5$, the confidence interval of the mean μ_{I_k} is defined as

$$C(Y_{I_k}) = \{\mu_{I_k}\colon \overline{Y}_{I_k} + t_{n_k-p,\frac{\alpha}{2}}\widehat{\sigma}_{I_k}\cdot\overline{C}_{I_k} \le \mu_{I_k} \le \overline{Y}_{I_k} + t_{n_{I_k}-p,1-\frac{\alpha}{2}}\widehat{\sigma}_{I_k}\cdot\overline{C}_{I_k}\},$$

where $E(\widehat{\sigma}_{I_k}^2) = \sigma_{I_k}^2$. Then $\widehat{\sigma}_{I_k}^2$ can be calculated as follows:

$$\widehat{\sigma}_{I_k}^2 = \frac{1}{n_{I_k}-P}\widehat{\varepsilon}_{I_k}^\top\widehat{\varepsilon}_{I_k} = \frac{1}{n_{I_k}-P}Y_{I_k}^\top(I_{n_{I_k}\times n_{I_k}} - H_{I_k})Y_{I_k},$$

where $\hat{\varepsilon}_{I_k} = Y_{I_k} - \hat{Y}_{I_k} = (I_{n_{I_k}\times n_{I_k}} - H_{I_k})Y_{I_k}$.

For the full-rank submatrix $X_{I_k}^\top X_{I_k}$, we have

$$H_{I_k} = X_{I_k}(X_{I_k}^\top X_{I_k} + \lambda I_{n\times n})^{-1}X_{I_k}^\top,$$

where λ is the disturbance term and $I_{n\times n})$ is a n-order identity matrix.

Consequently, the shortest confidence interval length for μ_{I_k} can be expressed as

$$L(C(Y_{I_k})) = 2\widehat{\sigma}_{I_k}\cdot\bar{C}_{I_k}\cdot t_{n_{I_k}-p,1-\frac{\alpha}{2}}.$$

III. Laplace LIC criterion for optimal subset selection

In this section, we present concrete steps for the optimal subset selection, which are known as Laplace LIC criterion. This criterion can effectively remove redundant information and shorten the interval length.

Step 1. For indicator subset sequence $\{I_k\}_{k=1}^{K_n}$, we select the optimal indicator subset based on the shortest interval length of μ_{I_k}. The subset I_{opt}^1 can be expressed as

$$I_{opt}^1 = \arg\min_{I_k}\{\widehat{\sigma}_{I_k}\cdot\bar{C}_{I_k}\cdot t_{n_{I_k-1},1-\frac{\alpha}{2}}\},$$

where $\widehat{\sigma}_{I_k}$, $\bar{C}_{I_k}$ and $t_{n_{I_k-1},1-\frac{\alpha}{2}}$ are obtained from $L(C(Y_{I_k})) = 2\widehat{\sigma}_{I_k}\cdot\bar{C}_{I_k}\cdot t_{n_{I_k-1},1-\frac{\alpha}{2}}$.

Step 2. The least squares of β based on Q_{I_k} and the variance of $\hat{\beta}_{Ik}$ are given by

$$\widehat{\beta}_{I_k} = (X_{I_k}^T X_{I_k})^{-1}X_{I_k}^T Y_{I_k}, \mathrm{var}(\widehat{\beta}_{I_k}) = \widehat{\sigma}_{I_k}^2(X_{I_k}^T X_{I_k})^{-1},$$

where $E(\widehat{\sigma}_{I_k}^2) = \sigma_{I_k}^2$. By maximizing the determinant of the information matrix $X_{I_k}^T X_{I_k}$, the indicator subset I_{opt}^2 is computed as follows:

$$I_{opt}^2 = \arg\max_{I_k}|X_{I_k}^T X_{I_k}|.$$

Step 3. When leveraging the intersection of two datasets to estimate the mean μ, the resulting confidence interval exhibits reduced length relative to that obtained from each dataset. The final optimal subset can be obtained as follows:

$$I_{opt} = I_{opt}^1 \cap I_{opt}^2.$$

Through the aforementioned steps, the optimal subset $Q_{I_{opt}} = (Y_{I_{opt}}, X_{I_{opt}})$ is selected from all subsets $\{Q = (Y_{I_k}, X_{I_k})\}_{k=1}^{K_n}$. The shortest interval length of $\mu_{I_{opt}}$ can be expressed as

$$L(C(Y_{I_{opt}})) = 2\widehat{\sigma}_{I_{opt}}\cdot\bar{C}_{I_{opt}}\cdot t_{n_{I_{opt}-1},1-\frac{\alpha}{2}}.$$

IV. Numerical analysis

In this section,we present a few examples to demonstrate the performance of the Laplace LIC criterion in optimal subset selection. Under identical experimental conditions, we also analyze the performance of two alternative methods for subsets I_{opt}^1 and I_{opt}^2. We calculate the MSE values and MAE values of estimators $\hat{\mu}$ with respect to I_{opt}^1, I_{opt}^2, and I_{opt}. Through experimental comparisons of MAE and MSE values across different methods, the performance of Laplace LIC is evaluated. Notably, lower MAE and MSE values indicate higher predictive stability.

A. Preparatory work

For the subsets, I_{opt}^1, I_{opt}^2, and I_{opt}, the local estimators are computed individually using the following formulations:

$$\hat{\mu}_{I_{opt}^1} = X_{I_{opt}^1}\widehat{\beta}_{I_{opt}^1}, \hat{\mu}_{I_{opt}^2} = X_{I_{opt}^2}\widehat{\beta}_{I_{opt}^2}, \hat{\mu}_{I_{opt}} = X_{I_{opt}}\widehat{\beta}_{I_{opt}}.$$

Then, the MSE values of the local estimators are calculated as follows:

$$\mathrm{MSE}(\hat{\mu}_{I_{opt}^1}) = \frac{1}{n_{I_{opt}^1}}[(Y_{I_{opt}^1} - \widehat{Y}_{I_{opt}^1})^\top(Y_{I_{opt}^1} - \widehat{Y}_{I_{opt}^1})],$$

$$\mathrm{MSE}(\hat{\mu}_{I_{opt}^2}) = \frac{1}{n_{I_{opt}^2}}[(Y_{I_{opt}^2} - \widehat{Y}_{I_{opt}^2})^\top(Y_{I_{opt}^2} - \widehat{Y}_{I_{opt}^2})],$$

$$\mathrm{MSE}(\hat{\mu}_{I_{opt}}) = \frac{1}{n_{I_{opt}}}[(Y_{I_{opt}} - \widehat{Y}_{I_{opt}})^\top(Y_{I_{opt}} - \widehat{Y}_{I_{opt}})].$$

Besides, the MAE values of the local estimators are defined by the following equations,

$$\mathrm{MAE}(\hat{\mu}_{I_{opt}^1}) = |\overline{Y}_{I_k} - \hat{\mu}_{I_{opt}^1}|,$$

$$\mathrm{MAE}(\hat{\mu}_{I_{opt}^2}) = |\overline{Y}_{I_k} - \hat{\mu}_{I_{opt}^2}|,$$

$$\mathrm{MAE}(\hat{\mu}_{I_{opt}}) = |\overline{Y}_{I_k} - \hat{\mu}_{I_{opt}}|.$$

B. Numerical simulation

It is assumed that the error term follows a Laplace distribution. The experimental dataset is generated from five distinct distributions: Geometric, Beta, Cauchy, Chi-square, and Uniform.

For dataset (Y, X), it is denoted as $Y = X\beta + \varepsilon, \varepsilon \sim Laplace(\mu_1, bI)$. X consists of (X_1, X_2), and Y is composed of (Y_1, Y_2). The notations involved are defined as follows:

$$\begin{aligned} &X_1 \in IR^{n_1 \times p}, \\ &X_2 \in IR^{n_2 \times p}, \\ &Y_1 = X_1\beta + \varepsilon_1, \quad n_1 = n - n_r, \\ &Y_2 = X_2\beta + \varepsilon_2, \quad n_2 = n_r. \end{aligned}$$

Specifically, $X_1 \sim N(0, 2)$, $X_2 = (X_{ij})$ follows other distribution in different cases.

Subsequently, X_2 is defined according to each of the following distributions:

(1) Geometric distribution: $X_2 \sim Geom(0.28)$,
(2) Beta distribution: $X_2 \sim Beta(2, 4)$,
(3) Cauchy distribution: $X_2 \sim Cauchy$,
(4) Chi-square distribution: $X_2 \sim \chi^2(3)$,
(5) Uniform distribution: $X_2 \sim Uniform(0, 1)$.

Besides, the error terms with respect to X_1 and X_2 are defined by $\varepsilon_1 \sim Laplace(0, \sigma 1)$, $\varepsilon_2 \sim Laplace(0, \sigma 2)$.

By varying the values of n and p, we test the stability of Laplace LIC criterion under different cases.

Example 1: Geometric distribution

Two experiments are demonstrated about the stability of the Laplace LIC criterion under geometric distribution conditions.

i: The impact of n on the stability of the LIC criterion.

The parameter configurations for the experiment are summarized below: $p = 8$, $K_n = 10$, $\alpha = 0.01$, $\sigma_1 = 2$, $\sigma_2 = 6$, $n_r = 50$, together with $n = 1000, 2000, 3000, 4000, 5000$.

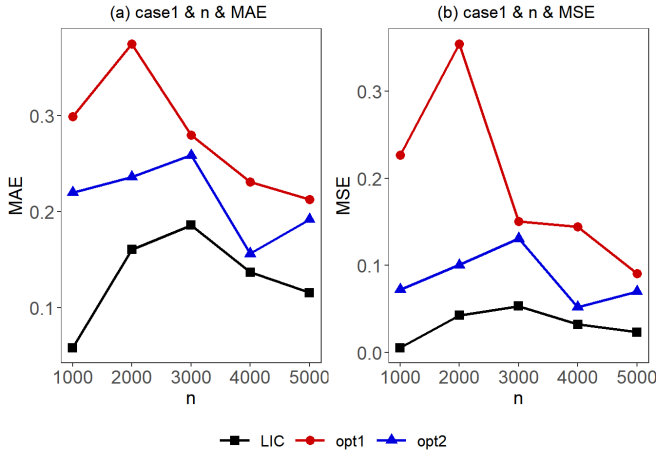

Fig. 1. Stability under geometric distribution when $p = 8$ and $n = (1000, 2000, 3000, 4000, 5000)$.

Fig. 1 demonstrates that under the geometric distribution, the LIC criterion exhibits superior stability in both MAE values and MSE values compared to the other two methods as the sample size increases. Specifically, when the sample size expands from 1000 to 3000, the MAE values exhibit relatively large fluctuations. However, within the range of 3000 to 5000, the MAE values gradually stabilize and show a downward trend. Meanwhile, the MSE values remain consistently low with minimal variability. These findings indicate that the LIC criterion achieves lower error margins and demonstrates high stability.

ii: The impact of p on the stability of the LIC criterion.

The parameter configurations for the experiment are summarized below: $n = 2000$, $K_n = 10$, $\alpha = 0.01$, $\sigma_1 = 2$, $\sigma_2 = 6$, $n_r = 50$, together with $p = 5, 6, 7, 8, 9$.

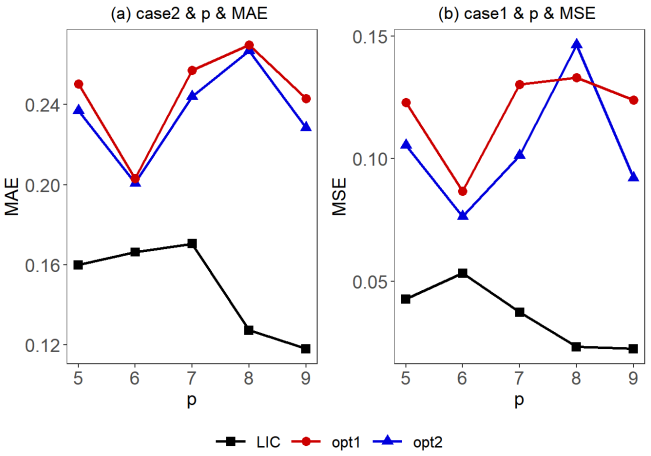

Fig. 2. Stability under geometric distribution when $n = 2000$ and $p = (5, 6, 7, 8, 9)$.

Fig. 2 demonstrates that under geometric distribution, both MAE values and MSE values for the other two methods exhibit significant fluctuations, whereas the LIC criterion maintains consistently low error values as p increases. Notably, when $p = 8$, the LIC criterion maintains low error levels compared to the other two approaches.

Example 2:Beta distribution

Two experiments are demonstrated about the stability of the Laplace LIC criterion under Beta distribution conditions.

i: The impact of n on the stability of the LIC criterion.

The parameter configurations for the experiment are summarized below: $p = 8$, $K_n = 10$, $\alpha = 0.01$, $\sigma_1 = 4$, $\sigma_2 = 6$, $n_r = 50$, together with $n = 1000, 2000, 3000, 4000, 5000$.

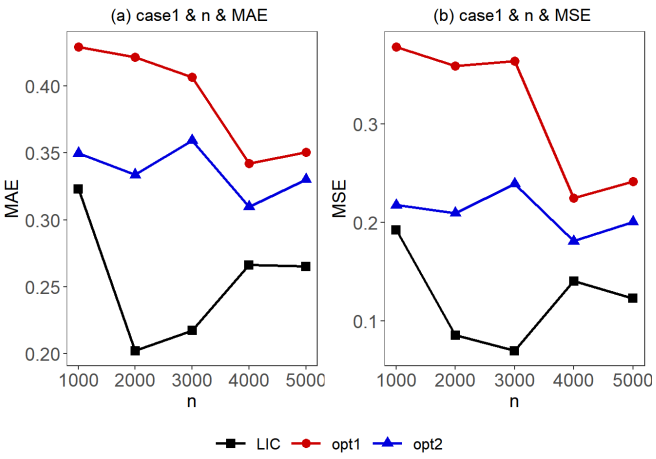

Fig. 3. Stability under beta distribution when $p = 8$ and $n = (1000, 2000, 3000, 4000, 5000)$.

Fig. 3 shows that under the Beta distribution, as the sample size increases, the MAE values and MSE values of the LIC criterion exhibit minor fluctuations while remaining consistently lower than the other two methods. These findings

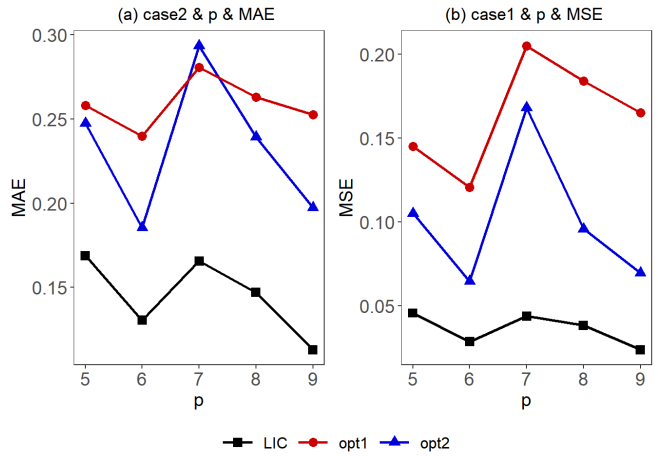

Fig. 4. Stability under beta distribution when $n = 2000$ and $p = (5, 6, 7, 8, 9)$.

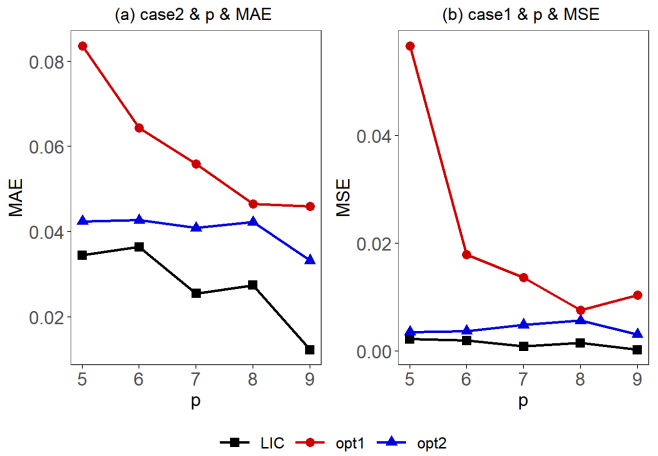

Fig. 6. Stability under cauchy distribution when $n = 2000$ and $p = (5, 6, 7, 8, 9)$.

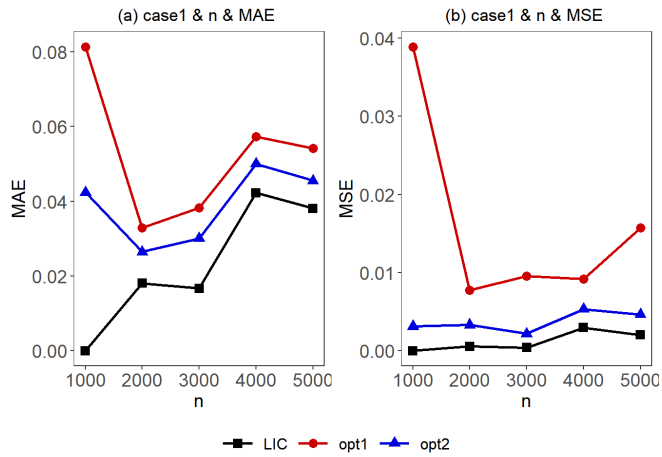

Fig. 5. Stability under cauchy distribution when $p = 8$ and $n = (1000, 2000, 3000, 4000, 5000)$.

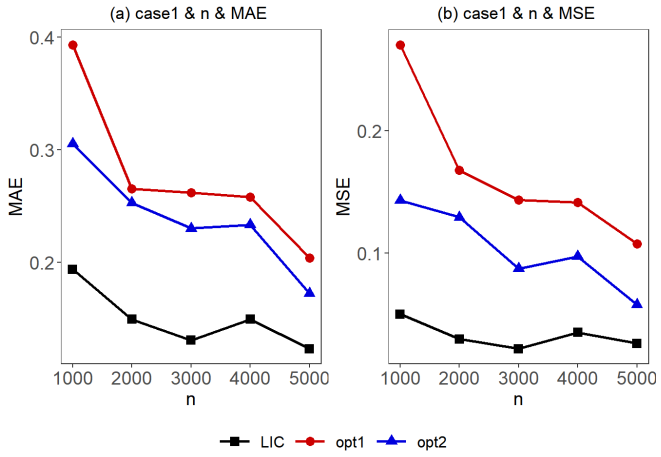

Fig. 7. Stability under chi-square distribution when $p = 8$ and $n = (1000, 2000, 3000, 4000, 5000)$.

indicate that the LIC criterion achieves smaller errors and superior performance.

ii: The impact of p on the stability of the LIC criterion.

The parameter configurations for the experiment are summarized below: $n = 2000$, $K_n = 10$,$\alpha = 0.01$, $\sigma_1 = 2$, $\sigma_2 = 8$, $n_r = 50$, together with $p = 5, 6, 7, 8, 9$.

Fig. 4 illustrates that under the Beta distribution, as the parameter p increases from 5 to 8, the MAE values and MSE values derived from the LIC criterion exhibit minimal fluctuations. In contrast, the other two methods show significant variability in MAE values and MSE values. Notably, when $p = 7$, the error values for the other two methods exhibit the largest fluctuations. These findings confirm that the LIC criterion offers superior stability and performance.

Example 3: Cauchy distribution

Two experiments are demonstrated about the stability of the Laplace LIC criterion under Cauchy distribution conditions.

i: The impact of n on the stability of the LIC criterion.

The parameter configurations for the experiment are summarized below: $p = 8$, $K_n = 10$, $\alpha = 0.01$, $\sigma_1 = 2$, $\sigma_2 = 6$, $n_r = 50$, together with $n = 1000, 2000, 3000, 4000, 5000$.

Fig. 5 demonstrates that under the Cauchy distribution, compared to the other two methods, the MAE and MSE values based on the LIC criterion are smaller. Notably, the MSE values derived from the LIC criterion exhibit less variability, demonstrating strong numerical stability.

ii: The impact of p on the stability of the LIC criterion.

The parameter configurations for the experiment are summarized below: $n = 2000$, $K_n = 10$, $\alpha = 0.01$, $\sigma_1 = 2$, $\sigma_2 = 6$, $n_r = 50$, together with $p = 5, 6, 7, 8, 9$.

Fig. 6 shows that under the Cauchy distribution, as the parameter p increases from 5 to 8, the LIC method exhibits a decreasing trend in both MAE values and MSE values, with minimal fluctuations. In contrast, the other two method displays significant variability in MAE and MSE values. These results confirm that the LIC criterion offers superior stability and performance.

Example 4: Chi-square distribution

Two experiments are demonstrated about the stability of the Laplace LIC criterion under Chi-square distribution conditions.

i: The impact of n on the stability of the LIC criterion.

The parameter configurations for the experiment are summarized below: $p = 8$, $K_n = 10$, $\alpha = 0.01$, $\sigma_1 = 2$, $\sigma_2 = 6$, $n_r = 50$, together with $n = 1000, 2000, 3000, 4000, 5000$.

Fig. 7 shows that under the Chi-square distribution, as the sample size increases, the LIC criterion exhibits a decreasing trend in both MAE and MSE values, with high stability. In contrast, while the other two methods also show decreasing trends in MAE and MSE, their error magnitudes remain significantly larger. These results indicate that the LIC method is more effective and demonstrates superior stability.

ii: The impact of p on the stability of the LIC criterion.

The parameter configurations for the experiment are summarized below: $n = 3000$, $K_n = 10$, $\alpha = 0.01$, $\sigma_1 = 4$, $\sigma_2 = 6$, $n_r = 50$, together with $p = 5, 6, 7, 8, 9$.

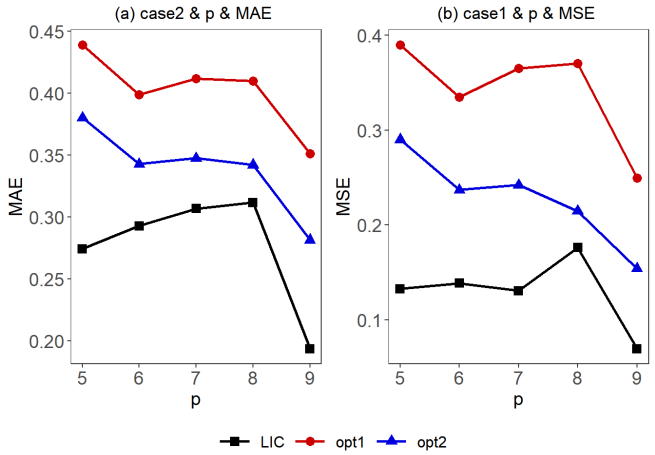

Fig. 8. Stability under chi-square distribution when $n = 3000$ and $p = (5, 6, 7, 8, 9)$.

Fig. 8 shows that under the Chi-square distribution, as the parameter p increases from 5 to 8, MAE values and MSE values for the LIC criterion remain consistently lower than those of the other methods.

Example 5: Uniform distribution

Two experiments are demonstrated about the stability of the Laplace LIC criterion under Uniform distribution conditions.

i: The impact of n on the stability of the LIC criterion.

The parameter configurations for the experiment are summarized below: $p = 8$, $K_n = 10$, $\alpha = 0.01$, $\sigma_1 = 4$, $\sigma_2 = 8$, $n_r = 50$, together with $n = 1000, 2000, 3000, 4000, 5000$.

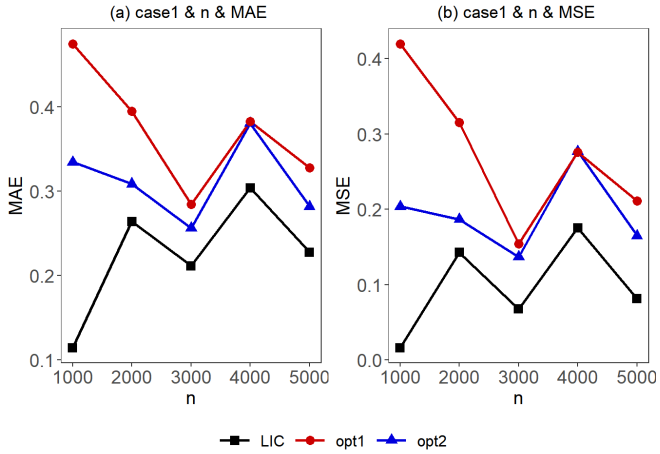

Fig. 9. Stability under uniform distribution when $p = 8$ and $n = (1000, 2000, 3000, 4000, 5000)$.

Fig. 9 shows that under the Uniform distribution, as the sample size increases, the MAE values and MSE values of all three methods exhibit fluctuations. However, the LIC criterion demonstrates a smaller range of fluctuations and lower error magnitudes compared to the others. These results indicate that the LIC criterion is both more effective and more stable.

ii: The impact of p on the stability of the LIC criterion.

The parameter configurations for the experiment are summarized below: $n = 2000$, $K_n = 10$, $\alpha = 0.01$, $\sigma_1 = 2$, $\sigma_2 = 6$, $n_r = 50$, together with $p = 5, 6, 7, 8, 9$.

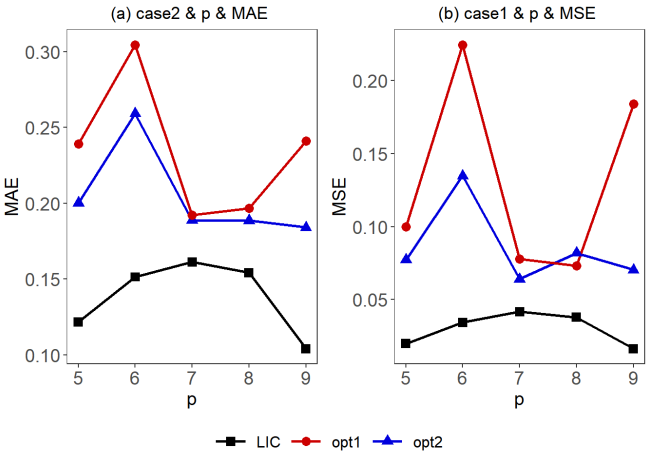

Fig. 10. Stability under uniform distribution when $n = 2000$ and $p = (5, 6, 7, 8, 9)$.

Fig. 10 shows that under the Uniform distribution, as the parameter p increases from 5 to 8, MAE vaues and MSE values for the LIC criterion exhibit minor fluctuations. In contrast, the other two methods show significantly larger fluctuations and higher error magnitudes.

III. Summary of the experiments.

Through numerical experiments, we analyze the stability and sensitivity of data under different distributions. The results demonstrate that the LIC criterion exhibits superior stability, effectively reducing errors and improving data reliability. Notably, as the data scale increases, the LIC criterion maintains a consistently lower error level.

V. Conclusion

This paper investigates the theoretical properties of the Laplace regression model, providing practical guidelines for redundant data processing. Specially, a comparative analysis is conducted about the Laplace LIC criterion. The results demonstrate that the Laplace LIC criterion demonstrates outstanding performance in optimal subset selection, exhibiting superior stability and sensitivity across diverse distributions.

Data Availability

We have implemented this criterion in the R package LLIC. Please visit the following website: https://cran.r-project.org/web/packages/LLIC/

References

[1] Q. Wang, G. B. Guo, G. Q. Qian, X. J. Jiang, "Distributed online expectation-maximization algorithm for Poisson mixture model," *Applied Mathematical Modelling*, vol. 124, no. 1, pp. 734–748, 2023.

[2] G. Guo, "Parallel statistical computing for statistical inference," *Journal of Statistical Theory and Practice*, vol. 6, pp. 536–565, 2012.

[3] J. Lederer, "Fundamentals of High-Dimension Statistics," in *Switzerland: Springer Nature Switzerland AG*, 2020. DOI: https://doi.org/10.1007/978-3-030-73792-4.
[4] M. Bottai, J. Zhang. "Laplace regression with censored data," Biom J 52:487–503, 2010.
[5] A. Yazdani, H. Zeraati, M. Yaseri, et al. "Laplace regression with clustered censored data," Comput Stat 37, 1041–1068, 2022.
[6] G. Guo, M. Yu, and G. Qian, "ORKM: Online regularized K-means clustering for online multi-view data," *Information Sciences*, vol. 680, p. 121133, 2024.
[7] G. Guo, H. Song, and L. Zhu, "The COR criterion for optimal subset selection in distributed estimation," *Statistics and Computing*, vol. 34, pp. 163–176, 2024.
[8] J. Fan and Y. Fan, "Distributed regression analysis under the restricted strong convexity condition," *Journal of the Royal Statistical Society: Series B (Statisticcal Methodology)*, vol. 80, no. 1, pp. 1–28, 2018.
[9] S. Chatterjee and S. N. Lahiri, "Bootstrapping L2 statistics in linear models with many covariates," *The Annals of Statistics*, vol. 39, no. 5, pp. 2442–2466, 2011.
[10] G. Guo, G. Qian, L. Zhu, "A scalable quasi-Newton estimation algorithm for dynamic generalized linear model," *Journal of Nonpara metric Statistics*, vol. 34, no. 4, pp. 917–939, 2022.
[11] M. Lin, N. Lin, and R. Chen, "Distributed robust regression with convex composite loss functions," *Journal of the American Statistical Association*, vol. 115, no. 529, pp. 900–913, 2020.
[12] T. Schifeling, Y. X. Wang, C. Sabatti, and E. J. Candes, "Distributed statistical estimation with sparse Gaussian graphical models," *Journal of Computational and Graphical Statistics*, vol. 25, no. 2, pp. 369–389, 2016.