

On almost unbiased ridge logistic estimator for the logistic regression model

Jibo Wu^{*†} and Yasin Asar[‡]

Abstract

Schaefer et al. [15] proposed a ridge logistic estimator in logistic regression model. In this paper a new estimator based on the ridge logistic estimator is introduced in logistic regression model and we call it as almost unbiased ridge logistic estimator. The performance of the new estimator over the ridge logistic estimator and the maximum likelihood estimator in scalar mean squared error criterion is investigated. We also present a numerical example and a simulation study to illustrate the theoretical results.

Keywords: Almost unbiased ridge logistic estimator, Ridge logistic estimator, Logistic regression model, Scalar mean squared error

2000 AMS Classification: 62J07

Received : 12.12.2014 *Accepted :* 04.06.2015 *Doi :* 10.15672/HJMS.20156911030

1. Introduction

In this paper we consider the estimation of Euclidean parameters $\beta \in R^p$ in logistic regression model based on the dependent variable y_i is $Be(\pi_i)$. The parameters π_i relate to β and x_1, x_2, \dots, x_n with the following value:

$$(1.1) \quad \pi_i = \frac{\exp(x_i\beta)}{1 + \exp(x_i\beta)}, \quad i = 1, 2, \dots, n$$

Usually the parameters of the model should be estimated using the maximum likelihood (ML) way by applying the following iterative weighted least square (IWLS) algorithm:

$$(1.2) \quad \hat{\beta}_{ML} = (X'WX)^{-1}X'\hat{W}\hat{Z}$$

where \hat{Z} is a vector with i th element equals $\log(\hat{\pi}_i) + \frac{y_i - \hat{\pi}_i}{\hat{\pi}_i(1 - \hat{\pi}_i)}$ and $W = \text{diag}(\frac{\hat{\pi}_i}{1 - \hat{\pi}_i})$.

^{*}Key Laboratory of Group & Graph Theories and Applications, Chongqing University of Arts and Sciences, Chongqing 402160, China, Email: linfen52@126.com

[†]Corresponding Author.

[‡]Department of Mathematics-Computer Science, Necmettin Erbakan University Konya, Turkey Email: yasar@konya.edu.tr, yasinasar@hotmail.com

Since ML estimation does not require any restriction on the characteristics of the independent variables, Maximum likelihood (ML) is the preferred estimation way in logistic regression. However, the ML estimator can be affected seriously by the presence of collinearity. It is known that ML parameter estimates have large variances in cases of multicollinearity. Many methods have proposed to combat this problem in linear regression model, such as the ridge estimator by Hoerl and Kennard [5], Liu estimator by Liu [10].

Schaefer et al. [15] use the ridge method to overcome the multicollinearity in logistic regression model and propose a ridge logistic estimator. Mansson and Shukur [13], Kibria et al. [12] proposed many methods to estimate the ridge parameter in ridge logistic estimator. Inan and Erdogan [9] proposed a Liu-type logistic estimator to overcome multicollinearity in logistic regression model.

Though the ridge logistic estimator proposed by Schaefer et al. [15] can overcome multicollinearity, however, this estimator has big bias. In this paper, we propose a new estimator which can be used not only overcome multicollinearity, but also can reduce the bias of the ridge estimator. We also discuss the statistical properties of the new estimator.

2. The almost unbiased ridge logistic estimator

The ridge logistic estimator (RLE) in the logistic regression model presented by Schaefer et al. [15] is denoted as follows:

$$(2.1) \quad \hat{\beta}_{RLE}(k) = (X'\hat{W}X + kI)^{-1}X'\hat{W}\hat{Z}, \quad k > 0$$

It is easy to obtain that:

$$\begin{aligned} Bias(\hat{\beta}_{RLE}(k)) &= E(\hat{\beta}_{RLE}(k)) - \beta \\ &= (X'\hat{W}X + kI)^{-1}X'\hat{W}X\beta - \beta \\ &= [(X'\hat{W}X + kI)^{-1}X'\hat{W}X - I]\beta \\ &= (X'\hat{W}X + kI)^{-1}[X'\hat{W}X - (X'\hat{W}X + kI)]\beta \\ (2.2) \quad &= -k(X'\hat{W}X + kI)^{-1}\beta \end{aligned}$$

and

$$(2.3) \quad Cov(\hat{\beta}_{RLE}(k)) = (X'\hat{W}X + kI)^{-1}X'\hat{W}X(X'\hat{W}X + kI)^{-1}$$

In linear regression model, many authors have studied the almost unbiased estimator, such as Kadiyala [11], Akdeniz and Kaciranlar [1] and Xu and Yang [16, 17].

To obtain the almost unbiased ridge logistic estimator, we firstly list the following definitions.

Definition 2.1. [16, 17] Suppose $\hat{\beta}$ is a biased estimator of parameter vector β , and if the bias vector of $\hat{\beta}$ is given by $Bias(\hat{\beta}) = E(\hat{\beta}) - \beta = R\beta$, which shows that $E(\hat{\beta} - R\beta) = \beta$, then we call the estimator $\tilde{\beta} = \hat{\beta} - R\hat{\beta} = (I - R)\hat{\beta}$ is the almost unbiased estimator based on the biased estimator $\hat{\beta}$.

Now, we are ready to derive the almost unbiased ridge logistic estimator based on the RLE. Since: $Bias(\hat{\beta}_{RLE}(k)) = (X'\hat{W}X + kI)^{-1}X'\hat{W}X\beta - \beta$, we have

$$\begin{aligned}
\hat{\beta}_{AURLE}(k) &= [I - ((X'\hat{W}X + kI)^{-1}X'\hat{W}X - I)]\hat{\beta}_{RLE}(k) \\
&= [2I - (X'\hat{W}X + kI)^{-1}X'\hat{W}X]\hat{\beta}_{RLE}(k) \\
&= [2I - (X'\hat{W}X + kI)^{-1}X'\hat{W}X](X'\hat{W}X + kI)^{-1}X'\hat{W}\hat{Z} \\
&= [I + (X'\hat{W}X + kI)^{-1}X'\hat{W}X](X'\hat{W}X + kI)^{-1}X'\hat{W}X\hat{\beta}_{ML} \\
&= [I + k(X'\hat{W}X + kI)^{-1}][I - k(X'\hat{W}X + kI)^{-1}]\hat{\beta}_{ML} \\
(2.4) \quad &= [I - k^2(X'\hat{W}X + kI)^{-2}]\hat{\beta}_{ML}
\end{aligned}$$

In the next section, we will discuss the properties of the new estimator.

For the convenience of the following discussions, let $\alpha = Q'\beta$, $\Lambda = diag(\lambda_1, \dots, \lambda_p) = Q'(X'\hat{W}X)Q$, where $\lambda_1 \geq \dots \geq \lambda_p > 0$ are the ordered eigenvalues of $X'\hat{W}X$.

3. The performance of the new estimator

The new estimator is proposed to reduce the bias of the ridge logistic estimator (RLE). So now we compare the new estimator with the RLE.

3.1. Theorem. *In logistic regression model we have*

$$\|Bias(\hat{\beta}_{AURLE}(k))\|^2 < \|Bias(\hat{\beta}_{RLE}(k))\|^2 \text{ for } k > 0.$$

Next we discuss the superiority of the new estimator in the scalar mean squared error (MSE) sense. Firstly we give its definition. Let $\hat{\beta}$ be an estimator of β , then the scalar mean squared error is defined as follows:

$$(3.1) \quad MSE(\hat{\beta}) = E(\hat{\beta} - \beta)'(\hat{\beta} - \beta) = tr\{Cov(\hat{\beta})\} + Bias(\hat{\beta})'Bias(\hat{\beta})$$

3.2. Theorem. *A sufficient of the new estimator superior to the RLE by the MSE criterion in logistic regression model is*

$$k > \frac{3 - \lambda_i\alpha_i^2 + \sqrt{(3 + \lambda_i\alpha_i^2)^2 + 4\lambda_i\alpha_i^2}}{4\alpha_i^2}$$

for all $i = 1, \dots, p$.

3.3. Theorem. *The new estimator is superior to the maximum likelihood (ML) estimator in logistic regression model for $k > 0$ if $1 - \lambda_i\alpha_i^2 > 0$ for all $i = 1, \dots, p$ and for $k < \frac{2\lambda_i + \lambda_i\sqrt{2(1 + \alpha_i^2\lambda_i)}}{\alpha_i^2\lambda_i - 1}$ if $1 - \lambda_i\alpha_i^2 < 0$ for some i .*

4. The selection of ridge parameter k

In this section we consider that the ridge parameter which is obtained by using the ridge parameter introduced in the previous section and the ridge parameters proposed by Hoerl and Kennard [5], Hoerl et al. [6], Batah et al. [3], Lawless and Wang [7] and Khurana et al [4].

The ridge parameter corresponding to Eq. (7.2) is

$$k_{NEW} = \frac{p}{\sum_{i=1}^p [\alpha_i^2 / [1 + (1 + \lambda_i * \alpha_i^2)^{1/2}]]}$$

Second, the Hoerl and Kennard [5] ridge parameter is defined as

$$k_{HK} = \frac{\hat{\sigma}^2}{\max \alpha_{iML}^2}$$

Table 1. Estimated quadratic bias with $\gamma = 0.9$

k	NEW	HK	HKB	LW	LS
RLE $\times 10^{-2}$	3.8401	0.5730	2.9811	0.0000	5.0532
AURLE $\times 10^{-2}$	0.0023	0.0000	0.0014	0.0000	0.0039

Table 2. Estimated quadratic bias with $\gamma = 0.95$

k	NEW	HK	HKB	LW	LS
RLE	0.1093	0.0020	0.0100	0.0000	0.1627
AURLE	0.0002	0.0006	0.0001	0.0000	0.0004

Table 3. Estimated quadratic bias with $\gamma = 0.99$

k	NEW	HK	HKB	LW	LS
RLE	0.9889	0.2645	0.9696	0.0000	1.3761
AURLE	0.1230	0.0090	0.0118	0.0000	0.2368

Third, the Hoerl et al. [6] ridge parameter is defined as

$$k_{HKB} = \frac{p\hat{\sigma}^2}{\hat{\beta}'_{ML}\hat{\beta}_{ML}}$$

Fourth, the Lawless and Wang [7] ridge parameter is defined as

$$k_{LW} = \frac{p\hat{\sigma}^2}{\hat{\beta}'_{ML}X'WX\hat{\beta}_{ML}}$$

Fifth, the Lindley and Smith [8] ridge parameter is defined as

$$k_{LS} = \frac{(n-p)(p+2)}{n+2} \frac{\hat{\sigma}^2}{\hat{\beta}'_{ML}\hat{\beta}_{ML}}$$

5. Monte Carlo simulation

The main purpose of this article is to compare the MSE properties and bias of the ML, RLE and AURLE when the regressors are highly intercorrelated. Hence, the core factor varied in the design of the experiment is the degree of correlation γ between the regressors. Therefore, the following formula which enables us to vary the strength of the correlation is used to generate the explanatory variables:

$$(5.1) \quad x_{ij} = (1 - \gamma^2)^{1/2} z_{ij} + \gamma z_{ip}, \quad i = 1, \dots, n, \quad j = 1, \dots, p$$

where z_{ij} are independent standard normal pseudo-random numbers, and γ is specified so that the correlation between any two explanatory variables is given by γ^2 .

Four different values of γ corresponding to 0.9, 0.95, 0.99 are considered and the sample size is equal to 50.

All simulation results are given in Tables 1-6.

From Tables 1-3, we can see that the new estimator has smaller quadratic bias than the RLE. When we see the estimated MSE of the new estimator and the RLE, we see that the new estimator is always superior to the RLE. The new estimator is superior to the RLE in the MSE criterion.

Table 4. Estimated MSE with $\gamma = 0.9$

k	0	NEW	HK	HKB	LW	LS
ML	0.0376	0.0376	0.0376	0.0376	0.0376	0.0376
RLE	0.0376	0.0709	0.0413	0.0629	0.0376	0.0823
AURLE	0.0376	0.0374	0.0376	0.0374	0.0376	0.0375

Table 5. Estimated MSE with $\gamma = 0.95$

k	0	NEW	HK	HKB	LW	LS
ML	0.0717	0.0717	0.0717	0.0717	0.0717	0.0717
RLE	0.0717	0.1645	0.0843	0.1558	0.0717	0.2146
AURLE	0.0717	0.0711	0.0714	0.0711	0.0717	0.0722

Table 6. Estimated MSE with $\gamma = 0.99$

k	0	NEW	HK	HKB	LW	LS
ML	0.3111	0.3111	0.3111	0.3111	0.3111	0.3111
RLE	0.3111	1.1128	0.4654	1.0948	0.3111	1.4769
AURLE	0.3111	0.2943	0.2955	0.3487	0.3111	0.4392

From the Tables, we also conclude that the new ridge parameter perform well.

6. Numerical example

In this section, we present a real data application in order to illustrate the benefits of the new estimator AURLE and satisfy the theoretical results. The data set is obtained from the official website of the Statistics Sweden (<http://www.scb.se/>) and it was also used in Asar and Genc [2] and a similar data set was used in Mansson et al. [4]. There are 271 observations which are the municipalities of Sweden in the data set. We fit a logistic regression model where the followings are the independent variables: the population (x_1), the number of unemployed people (x_2), the number of newly constructed buildings (x_3) and the number of bankrupt firms (x_4). We consider the net population change as the dependent variable such that it is coded as 1 if there is an increase in the population and 0 vice versa. We computed the bivariate correlations and observed that they are all greater than 0.90. The condition number being a measure of the degree of multicollinearity is computed as 38.3274 showing that there is severe multicollinearity problem with this data.

We provide the estimated theoretical MSE and coefficients of ML, RLE and AURLE for k_{NEW} , k_{HK} , k_{HKB} , k_{LW} and k_{LS} in Table 7.

It is observed from Table 7 that MSE of ML is the largest among all possible situations. The new estimator NEW works well with the estimator AURLE such that AURLE has a less MSE than RLE when NEW is used. Moreover, AURLE has better performance when HKB and LS are used. In Figure 1, we plot the MSE values of RLE and AURLE for changing values of the parameter k between zero and 1. It is seen from Figure 1 that AURLE has less MSE values in this interval. According to Theorem 3.2, for $k > 4.0608$, AURLE should have a less MSE than that of RLE. This result can be seen from Figure 2.

Table 7. The estimated theoretical MSEs and coefficients of estimators for different estimators of k

	k	β_1	β_2	β_3	β_4	SMSE
RLE	kNEW	2.3757	0.2549	1.1752	-2.7641	973.7117
	kHK	18.1501	-11.8274	3.6423	-9.0784	1157.0656
	kHKB	4.1873	-0.9625	1.9584	-4.1488	904.7912
	KLW	0.1674	0.1576	0.1342	0.0622	1076.8163
	kLS	8.3886	-4.2012	2.8882	-6.0829	826.1210
AURLE	kNEW	3.9688	-0.6164	2.0167	-4.3217	931.4176
	kHK	23.2398	-15.7957	3.8558	-10.4687	1642.0958
	kHKB	7.1375	-3.1442	2.9676	-5.9501	865.1787
	KLW	0.2679	0.2492	0.2012	0.0581	1074.8677
	kLS	13.6711	-8.3300	3.6388	-8.0437	936.7677
ML		25.3151	-17.4071	3.8669	-10.9661	1894.3979

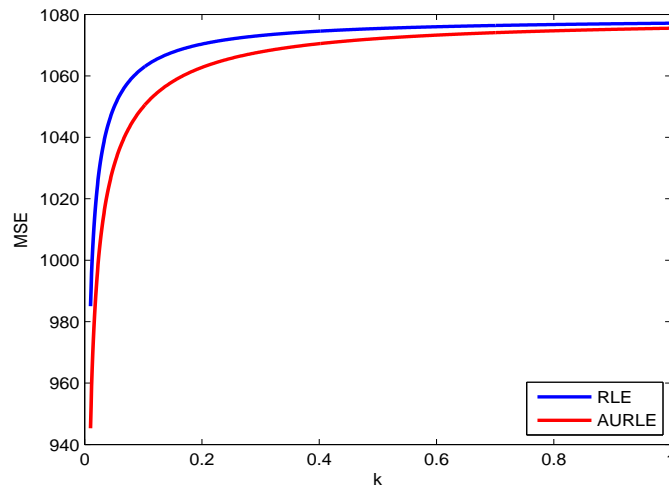


Figure 1. The estimated MSE of RLE and AURLE when $0 < k < 1$

Moreover, we plot the biases of the estimators to illustrate Theorem 3.1 in Figure 3. According to Figure 3, it is observed that the squared bias of AURLE is always less than that of RLE which coincides with Theorem 3.1.

Finally, Theorem 3.3 is also satisfied. Since $1 - \lambda_i \alpha_i^2 > 0$, AURLE has less MSE value than that of ML.

7. Conclusion

In this paper we propose a almost unbiased ridge logistic estimator based on the ridge logistic estimator and we also discuss the properties of the new estimator. The comparison results show that the new estimator has smaller quadratic bias the RLE, and under

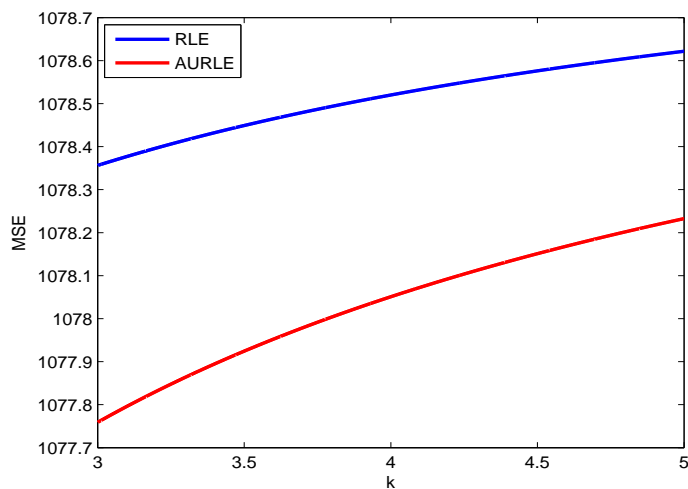


Figure 2. The estimated MSE of RLE and AURLE for satisfying Theorem 3.2

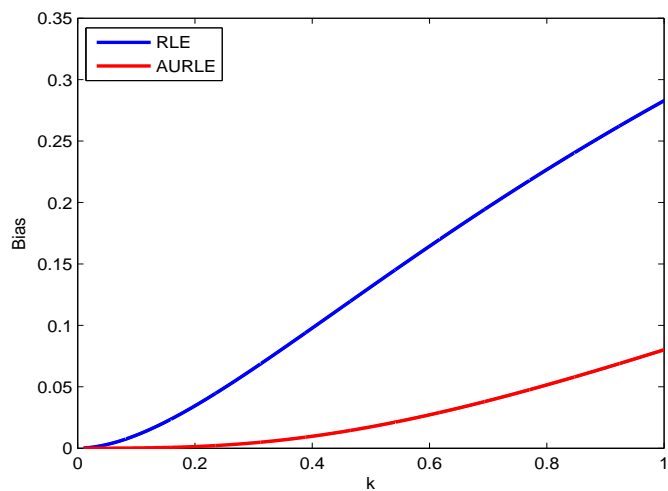


Figure 3. The biases of the estimator RLE and AURLE when $0 < k < 1$

certain conditions the new estimator is superior to the ML and RLE in the MSE sense.

Acknowledgements

The authors are thankful to the editor and the reviewers for their valuable comments and suggestions which have led to substantial improvement in the paper. This work was supported by the National Natural Science Foundation of China [No. 11501072],

the Natural Science Foundation Project of CQ CSTC [No. cstc2015jcyjA00001], and the Scientific and Technological Research Program of Chongqing Municipal Education Commission [No. KJ1501114]

References

- [1] Akdeniz, F. and Kaciranlar. S. *On the almost unbiased generalized Liu estimator and unbiased estimation of the Bias and MSE*, Comm. Statist. Theory Methods **24**, 1789-1797, 1995.
- [2] Asar, Y. and Genc, A. *New Shrinkage Parameters for the Liu-type Logistic Estimators*, Communications in Statistics-Simulation and Computation **45**(3), 1094-1103, 2016.
- [3] Batah, F.S.M. Gore, S. D. and Verma, V.R. *Effect of jackknifing on various ridge type estimators*, Model Assisted Statistics and Applications **3**, 201-210, 2008.
- [4] Khurana, M. Chaubey, Y.P. and Chandra, S. *Jackknifing the Ridge Regression Estimator: A Revisit*, Comm. Statist. Theory Methods **24**, 5249-5262, 2012.
- [5] Hoerl, A.E. and Kennard, R.W. *Ridge regression: Biased estimation for nonorthogonal problems*, Technometrics **12**, 55-67, 1970.
- [6] Hoerl, A.E. and Kennard, R.W. and Baldwin, K.F. *Ridge Regression: Some Simulations*, Comm. Statist. Theory Methods **4**, 105-123, 1975.
- [7] Lawless, J. F. and Wang, P. *A Simulation Study of Some Ridge and Other Regression Estimators*, Comm. Statist. Theory Methods **5**, 307-323, 1976.
- [8] Lindley, D. V. and Smith, A.F.M. *Bayes Estimate for The Linear Model (with discussion) part 1*, Journal of the Royal Statistical Society Ser B **34**, 1-41, 1972.
- [9] Inan, D. and Erdogan, B. E. *Liu-Type Logistic Estimator*, Communications in Statistics-Simulation and Computation **42**, 1578-1586, 2013.
- [10] Liu, K. *A new class of biased estimate in linear regression*, Comm. Statist. Theory Methods **22**, 393-402, 1993.
- [11] Kadiyala, K. *A class almost unbiased and efficient estimators of regression coefficients*, Econom. Lett **16**, 293-296, 1984.
- [12] Kibria, K. M. G. Mansson, K. and Shukur, G. *Performance of Some Logistic Ridge Regression Estimators*, Computational Economics **40**, 401-414, 2012.
- [13] Mansson, K. and Shukur, G. *On Ridge Parameters in Logistic Regression*, Comm. Statist. Theory Methods **40**, 3366-3381, 2011.
- [14] Mansson, K. Kibria, B. G. and Shukur, G. *On Liu estimators for the logist regression model*, Economic Modelling **29**(4), 1483-1488, 2012.
- [15] Schaefer, R.L. Roi, L.D. and Wolfe, R. A. *A ridge logistic estimator*, Comm. Statist. Theory Methods **13**, 99-113, 1984.
- [16] Xu, J. W. and Yang, H. *More on the bias and variance comparisons of the restricted almost unbiased estimators*, Comm. Statist. Theory Methods **40**, 4053-4064, 2011.
- [17] Xu, J. W. and Yang, H. *On the restricted almost unbiased estimators in linear regression*, Journal of Applied Statistics **38**, 605-617, 2011.

Appendix

3.1 Theorem

Proof. We have $Bias(\hat{\beta}_{RLE}(k)) = -k(X'\hat{W}X + kI)^{-1}\beta$ and

$$\begin{aligned} Bias(\hat{\beta}_{AURLE}(k)) &= [I - k^2(X'\hat{W}X + kI)^{-2}]\beta - \beta \\ (7.1) \qquad \qquad \qquad &= -k^2(X'\hat{W}X + kI)^{-2}\beta \end{aligned}$$

Thus we have

$$\begin{aligned} &\|Bias(\hat{\beta}_{RLE}(k))\|^2 - \|Bias(\hat{\beta}_{AURLE}(k))\|^2 \\ &= \beta'k^2(X'\hat{W}X + kI)^{-2}\beta - \beta'k^4(X'\hat{W}X + kI)^{-4}\beta \\ (7.2) \qquad \qquad \qquad &= \alpha'k^2(\Lambda + kI)^{-2}\alpha - \alpha'k^4(\Lambda + kI)^{-4}\alpha = \alpha'G\alpha \end{aligned}$$

where $G = k^2(\Lambda + kI)^{-2} - k^4(\Lambda + kI)^{-4} = diag(\frac{k^2\lambda_i(\lambda_i+2k)}{(\lambda_i+k)^4})$, thus for $k > 0$, $\alpha'G\alpha > 0$. The proof is completed. \square

3.2 Theorem

Proof. By (2.2)-(2.3) and the definition of SMSE, we have

$$\begin{aligned} MSE(\hat{\beta}_{RLE}(k)) &= tr\{Cov\hat{\beta}_{RLE}(k)\} + Bias(\hat{\beta}_{RLE}(k))'Bias(\hat{\beta}_{RLE}(k)) \\ &= tr\{(X'\hat{W}X + kI)^{-1}X'\hat{W}X(X'\hat{W}X + kI)^{-1}\} \\ &\quad + \alpha'k^2(\Lambda + kI)^{-2}\alpha \\ &= \sum_{i=1}^p \frac{\lambda_i}{(\lambda_i + k)^2} + \sum_{i=1}^p \frac{k^2\alpha_i^2}{(\lambda_i + k)^2} \\ (7.3) \qquad \qquad \qquad &= \sum_{i=1}^p \frac{\lambda_i + k^2\alpha_i^2}{(\lambda_i + k)^2} \end{aligned}$$

By (2.4), we can compute that:

$$(7.4) \qquad Cov\hat{\beta}_{AURLE}(k) = [I - k^2(X'\hat{W}X + kI)^{-2}](X'\hat{W}X)^{-1}[I - k^2(X'\hat{W}X + kI)^{-2}]$$

Then we get

$$\begin{aligned} &MSE(\hat{\beta}_{AURLE}(k)) \\ &= tr\{Cov\hat{\beta}_{AURLE}(k)\} + Bias(\hat{\beta}_{AURLE}(k))'Bias(\hat{\beta}_{AURLE}(k)) \\ &= tr\{[I - k^2(X'\hat{W}X + kI)^{-2}](X'\hat{W}X)^{-1}[I - k^2(X'\hat{W}X + kI)^{-2}]\} \\ &\quad + \alpha'k^4(\Lambda + kI)^{-4}\alpha \\ &= \sum_{i=1}^p (1 - \frac{k^2}{(\lambda_i + k)^2})^2 \frac{1}{\lambda_i} + \sum_{i=1}^p \frac{k^4\alpha_i^2}{(\lambda_i + k)^4} \\ (7.5) \qquad \qquad \qquad &= \sum_{i=1}^p \frac{(\lambda_i + 2k)^2\lambda_i + k^4\alpha_i^2}{(\lambda_i + k)^4} \end{aligned}$$

Now we consider the difference:

$$\begin{aligned} \Delta_1 &= MSE(\hat{\beta}_{RLE}(k)) - MSE(\hat{\beta}_{AURLE}(k)) \\ &= \sum_{i=1}^p \frac{\lambda_i + k^2\alpha_i^2}{(\lambda_i + k)^2} - \sum_{i=1}^p \frac{(\lambda_i + 2k)^2\lambda_i + k^4\alpha_i^2}{(\lambda_i + k)^4} \\ (7.6) \qquad \qquad \qquad &= \sum_{i=1}^p \frac{\lambda_i k [2k^2\alpha_i^2 + (\lambda_i\alpha_i^2 - 3)k - 2\lambda_i]}{(\lambda_i + k)^4} \end{aligned}$$

Δ_1 will be positive for $k > 0$ if and only if

$$(7.7) \quad 2k^2\alpha_i^2 + (\lambda_i\alpha_i^2 - 3)k - 2\lambda_i > 0$$

for all $i = 1, \dots, p$. The expression in (7.7) is a quadratic function of k which has two distinct roots

$$(7.8) \quad k_{1,2} = \frac{3 - \lambda_i\alpha_i^2 \pm \sqrt{(3 + \lambda_i\alpha_i^2)^2 + 4\lambda_i\alpha_i^2}}{4\alpha_i^2}$$

Though the root $\frac{3 - \lambda_i\alpha_i^2 - \sqrt{(3 + \lambda_i\alpha_i^2)^2 + 4\lambda_i\alpha_i^2}}{4\alpha_i^2}$ is negative. Thus when $k > 0$ and

$$k > \frac{3 - \lambda_i\alpha_i^2 + \sqrt{(3 + \lambda_i\alpha_i^2)^2 + 4\lambda_i\alpha_i^2}}{4\alpha_i^2}$$

for all $i = 1, \dots, p$, the new estimator is superior to the RLE by the MSE criterion in logistic regression model. \square

3.3 Theorem

Proof. It is easy to obtain that

$$(7.9) \quad MSE(\hat{\beta}_{ML}) = \sum_{i=1}^p \frac{1}{\lambda_i}$$

Now we study the following difference:

$$(7.10) \quad \begin{aligned} \Delta_2 &= MSE(\hat{\beta}_{ML}) - MSE(\hat{\beta}_{AURLE}(k)) \\ &= \sum_{i=1}^p \frac{1}{\lambda_i} - \sum_{i=1}^p \frac{(\lambda_i + 2k)^2\lambda_i + k^4\alpha_i^2}{(\lambda_i + k)^4} \\ &= k^2 \sum_{i=1}^p \frac{(1 - \alpha_i^2\lambda_i)k^2 + 4\lambda_i k + 2\lambda_i^2}{\lambda_i(\lambda_i + k)^4} \end{aligned}$$

Δ_2 will be positive if and only if

$$(7.11) \quad (1 - \alpha_i^2\lambda_i)k^2 + 4\lambda_i k + 2\lambda_i^2 > 0$$

Now we discuss $(1 - \alpha_i^2\lambda_i)k^2 + 4\lambda_i k + 2\lambda_i^2$.

(1) If $1 - \lambda_i\alpha_i^2 > 0$ for all $i = 1, \dots, p$, then $(1 - \alpha_i^2\lambda_i)k^2 + 4\lambda_i k + 2\lambda_i^2 > 0$.

(2) If $1 - \lambda_i\alpha_i^2 < 0$ for some $i = 1, \dots, p$, then using the method in Theorem 3.2, we have if $k < \frac{2\lambda_i + \lambda_i\sqrt{2(1 + \alpha_i^2\lambda_i)}}{\alpha_i^2\lambda_i - 1}$, $(1 - \alpha_i^2\lambda_i)k^2 + 4\lambda_i k + 2\lambda_i^2 > 0$.

This completes the proof of Theorem. \square