AGRICULTURAL JOURNAL

SINCE 2000
MEDWELL PUBLICATIONS

Statistical Analysis to Mammal Studies Based on Mammal Sleep Data

Liming Xie
North Dakota State University, North Dakota, USA

Key words: Lasso, Mammal sleep, adaptive Lasso, model selection, ASE, SBC (Schwarz Bayesian information), AIC (Akaike Information Criterion)

Corresponding Author:
Liming Xie
North Dakota State University, North Dakota, USA

Page No.: 97-106
Volume: 15, Issue 5, 2020
ISSN: 1816-9155
Agricultural Journal

Abstract: The researcher analyzes mammal sleep with 62 species in 1976 by using Lasso method (least absolute shrinkage and selection operator)that provides stability, higher selection variables, computational efficiency and higher prediction accuracy. the results of Average Parameter Estimate for using adaptive Lasso in SAS indicates that the position of slow wave and paradoxical sleep is account for 100%, overall danger index is 93%. The distributions of overall danger index and slow wave with paradoxical sleep as wee as gestation time from Refit model shows normal histogram for paradoxical sleep. In partition statement of "glmselect"procedure, ASE value (Average Square Error) of the validation from overall danger index is the minimum of all parameters in the selected model. On the other hand in selection steps for ASE, the adaptive Lasso method seems to have fewer than Lasso; for complicate and large data, elastic net can deal with more parameters than observations and combine one and a couple of groups that are consist of multiple variables by shrinking the coefficients of correlated variables toward each other.

INTRODUCTION

Lasso (least absolute shrinkage and selection operator) is one of popular regression statistical methods. It is mainly applied some advanced means such as variable selection or regularization to generate and interpret the data models in statistics, so that, deeply analyze and reveal the object data. It is introduced by Tibshirani[1] originally, Lasso was constructed with least squares models to display necessary amount on the estimators and the correlation each other, such as, the connection of ridge regression and coefficients. If so, users might to get the subset of predictors with minimum prediction error for a response variable. Lasso can perform zeroing out variable and shrinkage to improve the value of prediction that is it can select variables with less bias for variables that 'really matter', it can also allow much more parameters numbers than observations (but only include up to variables); discarding non-useful variables, etc. Many years ago, the studies for mammal sleep has been conducted. Giraffe's sleep at a zoo was tested with paradoxical sleep. To analyze behavioral sleep for mammals, many researchers conducted to regression study such as assessing 152 nights in 5 adults, 2 immatures and one juvenile giraffe at a zoological garden with PS (Paradoxical Sleep). The results showed that "ANOVA factor interval from 18-8 h"[2]. A scholar used logarithmic transformation to analyze correlation between sleep gestation and rapid eye movement for 79 mammal species[3]. Some limingxie2020@gmail.com potential covariates such as body weight, cage locations were screened with the sleep parameters and regarded a s normal distribution[4].

If two process modes and transition between different numbers of daily mammal sleep by wake were similar with mathematical models. For some mammals, the sleep/wake patterns of suprachiasmatic nucleus of the hypothalamus were simulated to be different values of the modulation parameter a that is between -1 and 1 at different periods[5]. Gamma distribution was used to analysis of the amplitude-frequency for spindle occurrences, the result showed that the mean frequency of the cluster of spindles shift from 11-13 Hz[6]. For female rats with asleep, two-way ANOVA was applied to analyze the distance and escape latency of group to test ovariectomized female Wistar rats if they are the normality[7]. However, these analyses bring into the following questions: they pick a model if a model selection is reasonable the selected model has been affected by outliers. Is the prediction more accurate? Do we deal with more parameters (that is p) than observations (that is n)? if so what statistical methods can we choose? Penalized regression method (adaptive Lasso and Elastic net) can perform the jobs that traditional selection methods such as backward, forward and stepwise selections cannot do more numbers of prediction variables than number of sample sizes and combining the group with multiple variables, etc.

MATERIALS AND METHODS

Allison and Cicchetti[8] published “Sleep in Mammals: Ecological and Constitutional Correlates” at A. A .A. Science. They thought that “slow-wave sleep was negative associated with a factor related to body size” and “paradoxical sleep was related to a factor with predatory danger,” based on data (See Supplemental Files S1).

It collected 62 mammals. It included the following variables: Species of animal, body weights (Body W) by kilogram; brain weight (Brain W) by grain, Slow Wave (SWS) with nondreaming sleep by (hrs/day), Paradoxical (PS) with dreaming sleep (hrs/day), Total Sleep (TS) with sum of slow wave and paradoxical sleep (hrs/day), maximum life span (LS, years), Gestation Time (GT, days), Predation Index (PI, point 1-5) at which point 1 denoted least likely to be preyed upon and point 5 denoted the most likely preyed upon, sleep exposure index (SEI, point 1-5) at which point 1 denoted the least exposed such as animal sleep in a well-protected den and point 5 denoted the most protected den, overall danger index (ODI, point 1-5) at which point 1 denoted the least danger from other animals and most danger from other animals according to the above two indicated and other information. Also, “-999” expresses missing values.

Two researchers pointed out that the species in the laboratory were not considered some factors such as environmental or ecological influences. For example, the definition of good sleeper and bad sleeper that in general, good sleeper was >8 h per day but for most mammals in the laboratory, they need have more time to be adaptive to the laboratory. Hence, some mammals did not meet the standardized time and sometimes, their sleep were not stable. For example, cat is good sleeper and rabbit is not good sleeper. So, the collection of their sleep time was subject to different sleep time requirement; some variables such as “slow wave sleep” was observed by the electroencephalogram that test behavioral and the acquiescence of autonomic nervous system; “paradoxical sleep” was defined as brief irregular activities of the extremities and facial muscle movement with the dreaming measured by a low-voltage electroencephalogram; “life-span” was calculated by their maximum amounts of time under natural environmental living without diseases and predator’s threat or other dangerous factors; “predation index” was rated by five-point scale that the probabilities to be preyed. For example if some slept in a burrow, den or well-protected position, point 1 were given; “overall danger index” was evaluated that the mammal’s danger to be preyed. For example if some species slept in the maximum exposure positions, then, they obtained a point 5, otherwise if the minimum exposure place, they got a value of 1. However, this analysis was rudimentary and did not have analyze by statistical methods in detail. Hence, i would like to analyze it by using Lasso selections to obtain more accurate estimate.

Statistical analysis: I take advantage of SAS 9.4 that manages analytics more readily to assess and estimate data characteristics. The statistical methods are Lasso selections and some traditional regression such as stepwise selections to analyze the mammal sleep data to further analyze variable correlation each other. In a linear mode, the response variable Y is modeled as a linear combination of the predictor variables, a_1, ..., a_p plus random noise that is $Y = \beta_0+\beta_1 a_{i1}+, ..., a_{ip}+\epsilon_I$.

Model selection: Statistical model selection joins and performs the predictive estimation for different models and selects a best model from among the alternatives. However, model selection is not easy to find an approximate best method of the truth and its accuracy of the model prediction. Moreover, it is not necessarily guaranteed the underlying truth. Because lasso can construct stable results for the data and more predictors than the sample size, this method is better than traditional selection such as forward, backward, stepwise selections. Also, it can shrink the regression coefficients, so, variable selection and coefficient estimation are worked same time.

In this study the model consists of Y and X_1-X_{10}: Y is species of animals; dependent X_4 that is PS; Body W, Brain W, SWS, PS, Life-Span, GT, PI, SEI, ODI are X_1-X_7, X_9, X_{10}, respectively.

Lasso selection: This is a specific selection with tuning value z that looks for the solution to the regression constrained minimizing objects:

$$\text{Argmin}_\beta \sum_{i=1}^{N}\left(a_i\left(b\beta\right)_i\right)^2 \text{ Subject to } \sum_{j=1}^{p}|\beta| \leq z$$

where, L_1 norm of the regression coefficients is the position of Lasso penalty. It simplifies the sum of their absolute values. In the regression coefficients, the shrinkage is used equally. Beforethe selection, each predictor variable could be standardized. Hence, the GLMSELECT procedure is the best candidate to perform this process and it could generate plots to track the selection process when using the coefficients by the same scale.

Adaptive lasso: Adaptive Lasso is a specific Lasso penalty that weights is used to each parameter to construct the Lasso constraint. These weights from adaptive Lasso manage shrinking more zero coefficients than shrinking the nonzero coefficients[9]:

$$\tilde{\beta} = \arg\min_\beta \sum_{i=1}^{N}\left(a_i\text{-}\left(b\beta\right)i\right)^2 \text{ subject to} \sum_{j=1}^{p}\left(Q_j\left[\beta_j\right]\right) \leq z$$

The Glmselect procedure performs the adaptive weights $\left(Q_j = 1/|\hat{\beta}|\right)$ by the ordinary least squares estimations of the regression coefficients. Also, it offers some options,for example, if more correlation variables and predictor variables are over the sample sizes, adaptive Lasso might mange stable regression coefficients. This is better than using coefficients of the ordinary least squares.

Elastic net: As i mention proceeding, Lasso cannot deal with the big data when more numbers of selected predictor variables are more than the sample sizes. Elastic net does not only solve this limitation, it does but also combine groups of correlate variables. For example, some objects for data share same specific pathways and build a group but you plan to identify these objects to become this group, then you might to try elastic net. It gets off these limitations that is more numbers of selected predictor variables than number of sample sizes and it combines all variables in formed group without ignoring any members. The following is optimal formula f or elastic net:

$$\tilde{\beta} = \arg\min_\beta \sum_{i=1}^{N}\left(a_i\text{-}\left(b\beta\right)i\right)^2 \text{ subject to}$$
$$\sum_{j=1}^{p}\left|\beta_j\right| \leq z_i \text{ and} \sum_{j=1}^{p}\beta_j^2 \leq z_2$$

Where the penalty of elastic net is the position of L_1 norm $\left(\sum_{j=1}^{p}\left|\beta_j^2\right|\right)$ and L_2 norm $\left(\sum_{j=1}^{p}\beta_j^2\right)$ (of the regression coeff icients. L_1 norm part does take variables selected by getting some coefficients as zero. L_2 norm norm performs group selection that helps shrink the coefficients of correlation variables each other. Therefore, we can write equation:

$$\tilde{\beta} = \arg\min_\beta\left(\sum_{i=1}^{N}\left(a_i\text{-}\left(b\beta\right)_i\right)^2 + \lambda_1\sum_{j=1}^{p}\left|\beta_j\right| + \lambda_2\sum_{j=1}^{p}\beta_j^2\right)$$

Here, both λ_1 and λ_2 are the tuning parameters.

RESULTS

As i mention preceding section, using model selection and model averagein Lasso, adaptive Lasso or elastic net methods can select prediction candidate models more accurately than other regression techniques. In forecast predictive models, adaptive Lasso method is reasonable to combine many variable selections to construct parsimonious predictive models. A model averaging for high-dimensional regression can participate into combine a couple of all of parameter selection or groups of multiple variables. It describes all of variables selected in the model corresponding to using the model. In the model, parameter estimates become the averages of the estimates for each sample. When a parameter is not chosen by linear model, the estimate value is defined as zero, the effect of shrinking the estimates of rarely picked parameters to be a zero. For example, in Table 1 average parameter Estimates using adaptive Lasso, i use glmselect procedure with "EffectSelectPct" choose of model average tables.

To understand the percentage of samples what the positionof each effect is in the selected model, I use the bar chart of the percentages for each parameter in the model shown in Fig. 1. It graphically describes their positions in the selected model iconically below.is the first one by 100% is the second one by 93%, last one is (about 8%).

To compute the average estimate for a parameter, we can partition the sum of the estimate values for that parameter in each sample by 800 samples. Those parameters of estimate values of zero in the model are not displayed in Fig. 2. But they are listed in another table (not appear in this paper due to the space limitation of the journal). In Fig. 2 we can see that the distributions of the estimates in refit model that are each parameter selected

Table 1: Average parameter estimate using adaptive Lasso technique

Parameters	Number non-zero	Non-zero Percentage	Average parameter estimate	Mean estimate	SD	25%	Estimate quantiles Median	75%
X_1	Intercept	100	100.00	2.876914	0.878777	2.515478	2.891615	3.249636
X_2	30	30.00	0.000221	0.000980	0	0	0.000124	
X_3	16	16.00	0.000056876	0.000646	0	0	0	
X_4	32	32.00	-0.024487	0.137463	0	0	0	
X_5	29	29.00	0.037510	0.132908	0	0	0.000200	
X_6	26	26.00	-0.007786	0.023335	0	0	0	
X_7	39	39.00	0.000562	0.000987	-0.000735	0	0	
X_8	25	25.00	0.146357	0.32020	0	0	0	
X_9	8	8.00	-0.007210	0.066569	0	0	0	
X_{10}	93	93.00	-0.558087	0.445068	-0.724999	-0.403263	-.291436	
X_1*X_2	*	13	13.00	-0.000000139	0.00000	0	0	0
X_3*X_4	*	100	100.00	-0.001028	0.000139	-0.001003	-0.001003	-.001003
X_6*X_{10}	*	29	29.00	0.002742	0.008105	0	0	0

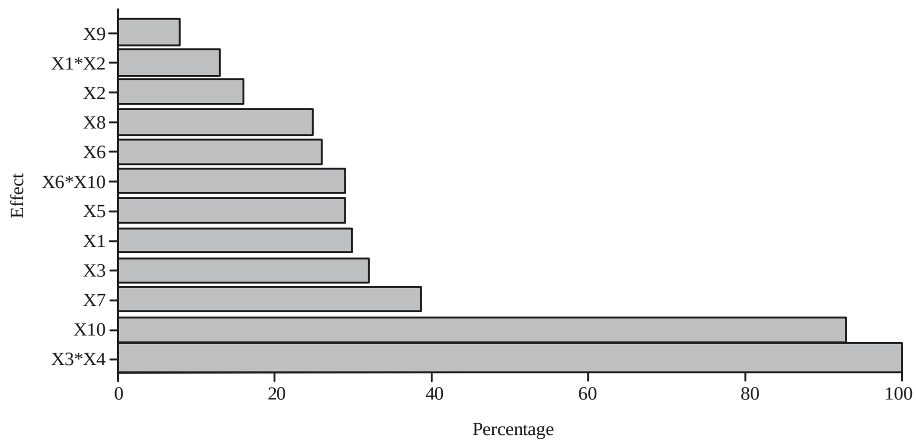

Fig. 1: The histogram for position of each effect in the model using adaptive Lasso (Effect selection percentage for X_4)

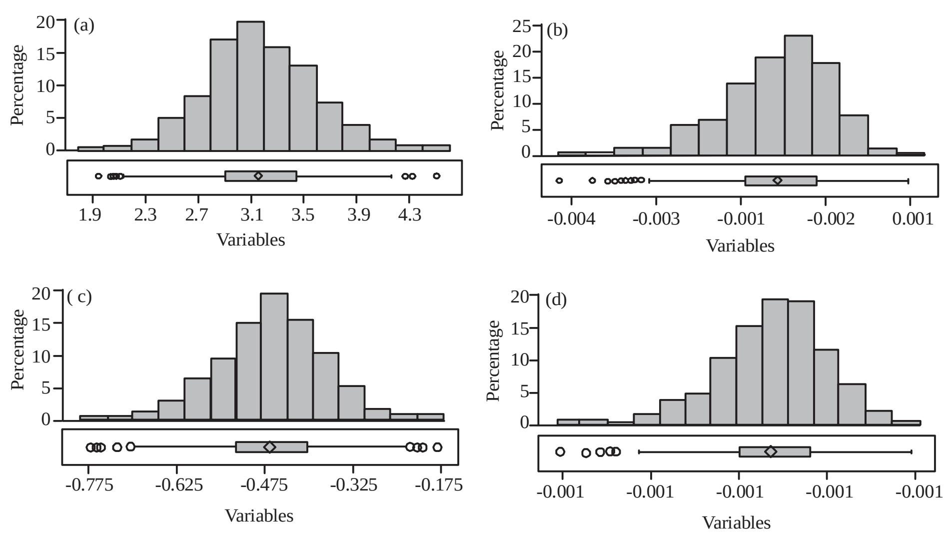

Fig. 2(a-d): The distributions of the estimates with model average refit using Lasso adaptive selection, choose is SBC. X_{10} is the closest to normal except intercept; Refit parameter estimate distribution for X_4 (a) Intercept (b) X_7 (c) X_{10} and X_3*X_4

Table 2: Information for Lasso SAS output

Dependent variables	Paradoxical sleep
Selection method	LASSO
Stop criterion	None
Choose criterion	Validation ASE
Effect Hierarchy enforced	None
Random number seed	686534000
Number of observations read	62
Number of observations used	62
Number of observation used for training	36
Number of observation used for validation	26
Dimensions	
Number of effects	56
Number of effects after splits	356

Table 3: Analysis of variance

Source	df	Sum of squares	Mean square	f-values
Model	7	4314928	616418	462767
Error	28	37.29680	1.33203	
Corrected total	35	4314966		
Root MSE	1.15414			
Dependent mean	-136.95556			
R-Square	1.0000			
Adj R^2	1.0000			
AIC	55.27399			
AICC	62.19707			
SBC	29.94214			
ASE(Train)	1.03602			
ASE(Validate)	1.38979			

Table 4: Lasso selection summary using Lasso

Dependent variable	Paradoxical sleep (X_4)
Selection method	LASSO
Stop criterion	None
Choose criterion	SBC
Effect Hierarchy Enforced	None
Random number seed	686534000
Number of observations read	44
Number of observations used	44
Number of observation read for test data	18
Number of observation used for test data	18
Dimensions	
Number of effects	10
Number of parameters	10

in the refit model. Because almost each distribution is approximately normal and large number of samples are used (sample size is 800). X_7, X_{10} and X_3* X_4 display the range between the fifth and 95th percentages of each estimate around (-0.00224, 0.00016), (-0.61116, -0.27128), (-0.00100, -0.00100), respectively.

In the above Table 2 construct a linear model by using “the Partition” statement that 40% of the data as validation data were reserved randomly and 60% as training data, so that, the prediction error of model selection could be estimated. This means that the training set is used to fit the models. For those large data, a validation set is the best method to tune a penalized regression technique. Those observations from the validation one would be used to produce a Lasso solution path and then finding a smallest ASE from the validation data. In addition, we can see out that 356 variables are selected due to classification of Y and as effects with specific levels and possible two-way interaction effects (unlisted). Also, in observation numbers, 26 observations are selected as validation data, 36 of remaining are regarded as training one (Table 3-5).

As you can see the Lasso selection summary, model selection is 42 total steps It shows that there are all the true 35 effects and X_{10} generates1.390 of the minimum validation ASE value in the step 7 with response variable. We can try to build a linear model in SAS programming such as GLMSELECT procedure to forecast the level of data. In this study i use first call for Lasso method and the other one is the adaptive Lasso. Both of them were TESTDATA = option for GLMSELECT procedure and the “CHOOSE” is SBC criterion for in the model statement. The partial SAS output are as follows:

For the criteria for AIC, AICC, SBC and Adjust R^2, they have different implications in the selected model: R^2 is commonly used to test accuracy in the model. It is a basic matrix to tell us what numbers of variance is in the model, its value reflects variable significance. For example, for 0.78 of R^2, 78% of the variation in the output variable is measuredby the input variables and adjusted R^2 is used to compute R^2 value of those variables that are increased into the model. So, adjust R^2 value is the statistic based on the independent variables in the model; AIC is an estimator of statistical models in the data. It provides a reference value for the model selection. It evaluates the relative information lost by a specific model that is if the information is lost in a model, then the model is higher

Table 5: The partial SAS output

Steps	Effect entered	Effect removed	Number effects in	SBC	ASE	Test ASE
0	Intercept		1	535.2239	175988.5830	112403.574
1	X_3		2	482.8579	24840.1650	40657.928
2	X_5		3	434.9205	15161.9390	31445.192
3	X_{10}		4	423.7731*	10798.7870	27318.165
4	X_9		5	426.5312	10549.8840	27359.058
5	X_8		6	429.1448	10272.9130	27699.398
6	X_7		7	432.1285	10087.7090	27960.554
7		X_9	6	427.5976	9917.9420	28346.309
8	X_6		7	431.3062	9900.9230	28386.1870
9		X_7	6	425.2497	9402.5920	30070.7370
10	X_1		7	428.0058	9185.4310	31653.1570
11	X_9		8	431.4157	9107.6280	32327.0090
12	X_2		9	433.3928	8741.1470	39119.2390
13	X_7		10	436.1181	8533.3080	71832.4010

*Optimal value of criterion selection stopped because all effects are in the final model

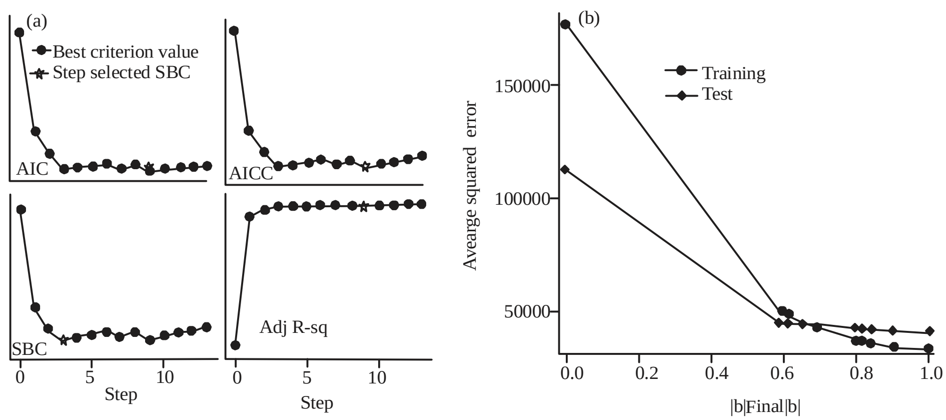

Fig. 3(a, b): The criterion plot for AIC, AICC, SBC and adjust R^2 selected in the model and b Progression of ASE by Role for X_4

quality; AICC is a kind information of AIC in statistical model for correction of small sample sizes; SBC is a very important estimator that selects best predict subsets in the regression models. Both of AIC and SBC have different goals. AIC is a good estimator for distance with the likelihood functions between fitting one and unknown one in the model. If the AIC value is lower, the model will be closed to the truth; SBC is a function estimator for testing if a model is true. When SBC is lower, the model is more likely to the real model. However, we should analyze them based on various assumptions approximations. In Fig. 3a for dependent X4, it has lower value of AIC and AICC at step 9, SBC is the minimum value at step 3, the best value of Adjust R-Square is at step 9. In Fig. 3b, we can see out that the amount of shrinking. In the regression coefficients, this shrinkage is decreased. But the model increase complexity, the ASE on the training data consistently dropping approximation to zero. Also, the prediction error on the test data is decreased by about 0.6 which is the point of minimum ASE in the vertical line of the plot. For test error, the decrease shows that the effects that join the model are important effects for the variation in the response variable before the vertical line. The subsequent increase suggests that the later effects explain the random noise in the training data. Hence, when reaching the minimum ASE value for the test data, the model is selected.

In Fig. 4 plots show that for using SBC criterion of model selection, Lasso and adaptive Lasso have the same set of predictor variables (X_3, X_2, X_{10}), although they are different solution paths. Also, the estimated coefficient values are near the same patterns.

In Table 6 the ASE of the test data form adaptive Lasso (27194) is little lower than one of Lasso (27318). Other values for adaptive Lasso are slightly smaller than corresponding ones for Lasso. Adaptive Lasso has a character to be distinguishing from big data sets. In GLMSELECT package, the former is also less steps than the latter (Fig. 4). Probably because the adaptive Lasso has a relatively higher penalization for zero coefficients and lower penalization for nonzero coefficients. So, it can decrease the estimation bias and advance variable selection accuracy, although Lasso has also an advantage with solution of difficult prediction problems (Table 7).

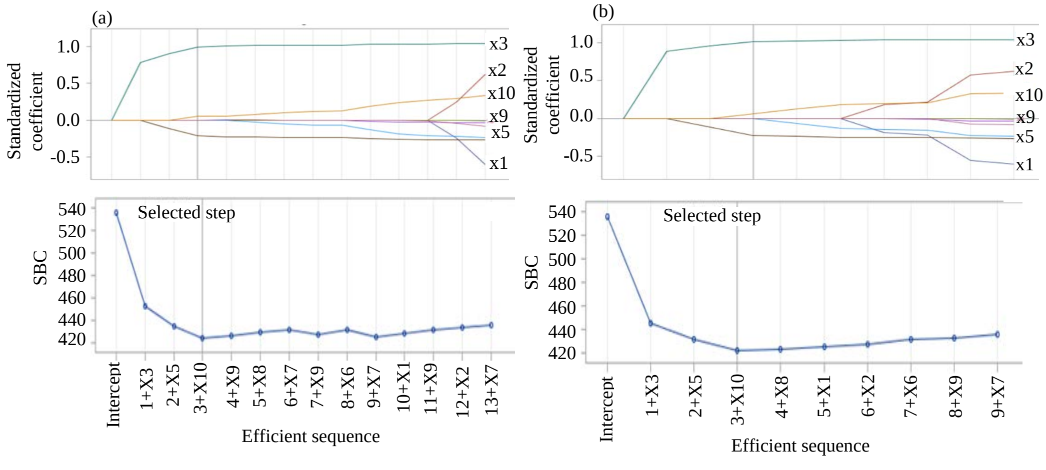

Fig. 4(a, b): Plot of coefficient progression for Lasso with adaptive Lasso (a) Lasso coefficient progression and (b) Adaptive Lasso coefficient progression

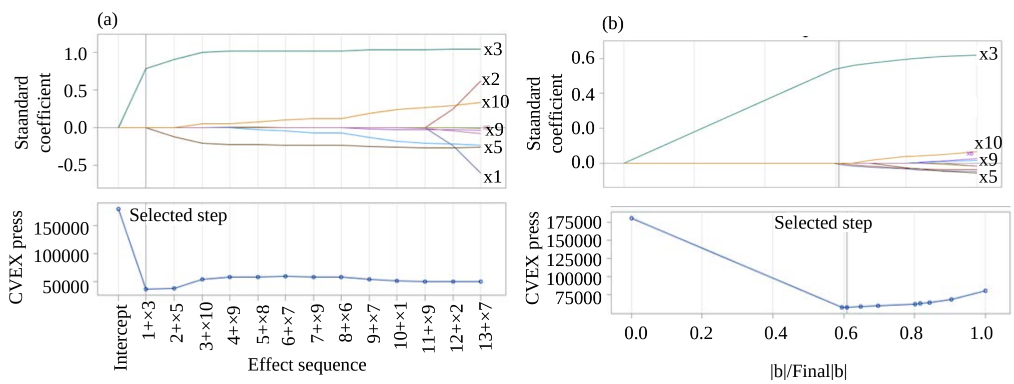

Fig. 5(a, b): Plot of coefficient progressions of Lasso and Elastic Net. X_3, X_6, X_{10} variables enter the model in the elastic net as a group before other candidates such as X_5 and X_9, etc. But it is not show up the group selection. Additionally, its solution path from elastic net looks more stable and smoother than Lasso path (a) Coefficient progression for X_4 and (b) Coefficient progression for X_4

Table 6: Fit statistics for Lasso and Adaptive Lasso

Selected models	Lasso	Adaptive Lasso
Root MSE	108.98929	107.58817
Dependent mean	-225.46591	-225.46591
R^2	0.9386	0.9402
Adj R^2	0.9340	0.9357
AIC	462.63632	461.49769
AICC	464.21527	463.07664
SBC	423.77308	422.63446
ASE (Train)	10799	10523
ASE (Test)	27318	27194

Table 7: Parameter estimates and fit statistics table for elastic net

Selected model			Root MSE	132.81225
Parameter estimates			Dependent mean	-225.46591
Parameter	df	Estimates	R-squared	0.9089
Intercept	1	-55.793540	Adj R-sq	0.9020
X_3	1	0.31710	AIC	480.03276
X_1*X_2	1	-0.00000396	AICC	481.61171
X_3*X_4	1	-0.000423	SBC	441.16952
			ASE (Train)	16036
			ASE (Test)	15025
			CVEX press	20296

Because the linear model of this case has the response variable, X_4 and X_1, X_2, X_3, X_5, …, X_{10} are the explanatory variable, it is very important to generate a better selection procedure identify a group such as letting and to get together. This is a key process to deal with a complicate and large data, especially for more parameters than observations. For Lasso, one limitation of variable selection is that the predict variables cannot be over sample size and it could limit groups of correlated variables. It takes only one variable for group and getting off remaining variables. However, elastic net method does not have those limitations for selected variable numbers and group selection numbers when generating the groups. I would like to try to use some techniques to explore solving more difficult problems. The following is the plots that apply elastic net method by using the tuning value of as 0.1, taking external cross validation to determine the tuning with cross validation for Lasso, it also cross validation technique as CVMETHOD choose. I would like to test if both have big difference. Lasso coefficient progressionb. Elastic Net coefficient progression.

The following table is parameter estimates of elastic net table with three predictors, X_3, X_1*, X_2*, X_3*, X_4. Here, X_1 X_2*(-0.000000396) and * (-0.000423) are so close to zero, ASE (Test) is 15025 that is significantly less than one of Lasso (27318) and Adaptive Lasso (27194).

DISCUSSION

Lasso technique as a new regression methodis involved penalizing the absolute values of the regression coefficients. It is a very important analysis approach to study and explore mammal research. Lasso selections including adaptive, elastic net and group Lasso are mature technique to help analyze and solve regression problems of comparative animal physiology. This study conforms some of the consequences based on mammal sleep data in 1976: Paradoxical sleep is correlation coefficient with slow wave Paradoxical sleep is also subject to predatory danger. In addition, this study tells us that body weight and brain weight are correlation with paradoxical sleep (Appendix 1 and 2).

CONCLUSION

I got another conclusion is that using adaptive Lasso and elastic net methods. They help deal with complicate and large data when more parameters than observations in the model. In prediction model, i used its stability, higher prediction accuracy, computational efficiency, higher selection methods of adaptive Lasso to generate GLMSECT procedure that performs model selections. It results in the effect of shrinking the estimates without zero value parameters. It guarantees higher accuracy for prediction models.

The GLMSECT technique is one of powerful model selection procedures. It can provide options and higher graphics to control selection by extensive customization. Also, if we need make partition of big data to training, validation, defining spline effects, selecting individual levels of classification effect, test sets, a couple of fit criterion such as AIC, AICC, SBC or k-fold cross validation to estimate prediction error, it will give powerful support.

ACKNOWLEDGEMENTS

I thank editors and three of anonymous reviewers. They gave constructive comments and suggestions.

APPENDIXES

Appendix 1: S1 the mammal sleep data from 1976 used in main analysis

Species	Body W	Brain W	SWS (h/day)	PS (h/day)	TS (h/day)	LS years	GT Days	PI	SEI	ODI
African elephant	6654	5712	-999	-999	3.30	38.6	645	3	5	3
African rat	1	6.6	6.3	2.0	8.30	4.50	42	3	1	3
Arctic fox	3.385	44.5	-999	-999	12.5	14.0	60	1	1	1
Arctic ground squirrel	0.920	5.70	-999	-999	16.5	-999	25	5	2	3
Asian elephant	2547	4603	2.1	1.8	3.90	69.0	624	3	5	4
Baboon	10.55	179.5	9.1	0.7	9.80	27.0	180	4	4	4
Big brown bat	0.023	.300	15.8	3.9	19.7	19.0	35	1	1	1
Brazilian tapir	160	169	5.2	1.0	6.20	30.4	392	4	5	4
Cat	3.30	25.6	10.9	3.6	14.5	28.0	63	1	2	1
Chimpanzee	52.16	440	8.3	1.4	9.7	50.0	230	1	1	1
Chinchilla	0.425	6.40	11.0	1.5	12.5	7.00	112	5	4	4
Cow	465	423	3.20	0.7	3.90	30.0	281	5	5	5
Desert hedgehog	0.550	2.40	7.6	2.7	10.3	-999	-999	2	1	2
Donkey	187.1	419	-999	-999	3.10	40.0	365	5	5	5
Eastern A. mole	0.075	1.20	6.3	2.1	8.40	3.50	42	1	1	1
Echidna	3.00	25.0	8.6	0.00	8.60	50.0	28	2	2	2
European hedgehog	0.785	3.50	6.6	4.1	10.7	6.00	42	2	2	2
Galago	0.200	5.00	9.5	1.2	10.7	10.4	120	2	2	2
Genet	1.41	17.5	4.8	1.3	6.10	34.0	-999	1	2	1
Giant armadillo	60.0	81.0	12.0	6.1	18.1	7.00	-999	1	1	1

Appendix 1: Continue

Species	Body W	Brain W	SWS (h/day)	PS (h/day)	TS (h/day)	LS years	GT Days	PI	SEI	ODI
Giraffe	529	680	-999	0.30	-999	28.0	400	5	5	5
Goat	27.66	115	3.3	0.50	3.80	20.0	148	5	5	5
Golden hamster	0.120	1.00	11.0	3.4	14.4	3.90	16.0	3	1	2
Gorilla	207	406	-999	-999	12.0	39.3	252	1	4	1
Gray seal	85.0	325	4.7	1.5	6.20	41.0	310	1	3	1
Gray wolf	36.33	119.5	-999	-999	13.0	16.2	63.0	1	1	1
Ground squirrel	0.101	4.00	10.4	3.4	13.8	9.00	28.0	5	1	3
Guinea pig	1.04	5.50	7.40	0.80	8.20	7.60	68.0	5	3	4
Horse	521	655	2.10	0.80	2.90	46.0	336	5	5	5
Jaguar	100	157	-999	-999	10.8	22.4	100	1	1	1
Kangaroo	35.0	56.0	-999	-999	-999	16.3	33.0	3	5	4
Lesser short-tailed shrew	0.005	.140	7.7	1.40	9.10	2.60	21.5	5	2	4
Little brown bat	0.010	0.25	17.9	2.00	19.9	24.0	50.0	1	1	1
Man	62.0	1320	6.10	1.90	8.00	100	267	1	1	1
Mole rat	0.122	3.00	8.20	2.40	10.6	-999	30.0	2	1	1
Mountain beaver	1.35	8.10	8.40	2.80	11.2	-999	45.0	3	1	3
Mouse	0.023	0.400	11.9	1.30	13.2	3.20	19.0	4	1	3
Musk shrew	0.048	0.330	10.8	2.00	12.8	2.00	30.0	4	1	3
N. American Opossum	1.70	6.30	13.8	5.60	19.4	5.00	12.0	2	1	1
Nine-banded Armadillo	3.50	10.8	14.3	3.10	17.4	6.50	120	2	1	1
Okapi	250	490	-999	1.00	-999	23.6	440	5	5	5
Owl monkey	0.480	15.5	15.2	1.80	17.0	12.0	140	2	2	2
Patas monkey	10.0	115	10.0	.900	10.9	20.2	170	4	4	4
Phalanger	1.62	11.4	11.9	1.80	13.7	13.0	17.0	2	1	2
Pig	192	180	6.50	1.90	8.40	27.0	115	4	4	4
Rabbit	2.50	12.1	7.50	0.900	8.40	18.0	31.0	5	5	5
Racoon	4.288	39.2	-999	-999	12.5	13.7	63.0	2	2	2
Rat	0.280	1.90	10.6	2.60	13.2	4.70	21.0	3	1	3
Red fox	4.235	50.4	7.40	2.40	9.80	9.80	52.0	1	1	1
Rhesus monkey	6.80	179	8.40	1.20	9.60	29.0	164	2	3	2
Rock hyrax (Hetero b)	0.750	12.3	5.70	0.900	6.60	7.00	225	2	2	2
Rock hyrax (Procavia)	3.60	21.0	4.90	0.500	5.40	6.00	225	3	2	3
Roe deer	14.83	98.2	-999	-999	2.60	17.0	150	5	5	5
Sheep	55.5	175	3.20	0.600	3.80	20.0	151	5	5	5
Slow loris	1.40	12.5	-999	-999	11.0	12.7	90.0	2	2	2
Star nose mole	0.060	1.00	8.10	2.20	10.3	3.50	-999	3	1	2
Tenrec	0.900	2.60	11.0	2.30	13.3	4.50	60.0	2	1	2
Tree hyrax	2.00	12.3	4.90	0.500	5.40	7.50	200	3	1	3
Tree shrew	0.104	2.50	13.2	2.60	15.8	2.30	46.0	3	2	2
Vervet	4.19	58.0	9.70	0.600	10.3	24.0	210	4	3	4
Water opossum	3.50	3.90	12.8	6.60	19.4	3.00	14.0	2	1	1
Yellow-bellied marmot	4.05	17.0	-999	-999	-999	13.0	38.0	3	1	1

Appendix 2: S2 Lasso selection summary

Steps	Effect entered	Effect removed	Number effects in	ASE validation ASE
0	Intercept	1	119660.157	214139.504
1	X_4_999	2	3.267	3.273
2	X_3	3	2.944	2.770
3	X_9	4	2.388	2.202
4	X_4_6.6	5	2.320	2.153
5	X_7	6	1.847	1.963
6	X_4_6.1	7	1.575	1.707
7	X_{10}	8	1.036	1.390*
8	X_5*X_4_999	9	0.961	6.421
9	X_4_4.1	10	0.868	19.925
10	X_4_0	11	0.678	69.878
11	X_4_3.9	12	0.603	95.712
12	X_4_3.6	13	0.392	189.714
13	X_3	12	0.323	237.965
14	X_4_2.6	13	0.311	244.668
15	X_4_3.1	14	0.305	247.054
16	X_5*X_4_0.6	15	0.292	251.499
17	X_4_2.8	16	0.253	258.313
18	X_7*X_4_1.4	17	0.207	270.884
19	X_4_1.9	18	0.190	277.082
20	X_4_1.3	19	0.176	284.200

Appendix 2: Continue

Steps	Effect entered	Effect removed	Number effects in	ASE validation ASE
21	*_0.8	20	0.126	309.601
22	X_5 *X_4_2.6	21	0.107	321.607
23	X_3 *X_4_1.4	22	0.106	322.049
24	X_4_1.2	23	0.103	323.994
25	X_4	24	0.102	324.412
26	X_4_0.3	25	0.102	338.034
27	X_4	24	0.065	338.034
28	X_4_2.4	25	0.057	336.192
29	X_7* X_4_0.6	26	0.057	335.958
30	X_4_0.9	27	0.042	325.076
31	X_4_1	28	0.037	322.751
32	X_6	29	0.035	323.450
33	X_4_1.8	30	0.035	323.680
34	X_8* X_4_1.5	31	0.027	334.412
35	X_7* X_4_2.4	32	0.020	353.610
36	X_4_0.7	33	0.015	350.392
37	X_1* X_4_1.5	34	0.013	350.782
38	X_4_2.4	33	0.009	342.219
39	X_1*X_4_2	34	0.007	332.143
40	X_6*X_{10}	35	0.005	335.019
41	X_6* X_{10}_2.6	34	0.003	337.522
42	X_{10}* X_4_2.6	35	0.001	346.241

*Optimal value of criterion selection stopped because the selected model is a perfect fit

REFERENCES

01. Tibshirani, R., 1996. Regression shrinkage and selection via. the lasso. J. Royal Stat. Soc.: Ser. B (Methodol.), 58: 267-288.
02. Tobler, I. and B. Schwierin, 1996. Behavioral sleep in the giraffe in a zoological garden. J. Sleep Res., 5: 21-32.
03. Gonfalone, A.A., 2016. Negative correlation between gestation and sleep durations in mammals. Open Access Anim. Physiol., 8: 1-7.
04. Golub, M.S. and C.E. Hogrefe, 2014. Sleep patterns in male juvenile monkeys are influenced by gestational iron deprivation and monoamine oxidase a genotype. Br. J. Nutr., 112: 1478-1483.
05. Phillips, A.J., B.D. Fulcher, P.A. Robinson and E.B. Klerman, 2013. Mammalian rest/activity patterns explained by physiologically based modeling. PLoS Comput. Biol., Vol. 9, No. 9. 10.1371/journal. pcbi.1003213.
06. Panas, D., U. Malinowska, T. Piotrowski, J. Zygierewicz and P. Suffczynski, 2013. Statistical analysis of sleep spindle occurrences. PLoS One, Vol. 8, No. 4.
07. Saadati, H., S. Esmaeili-Mahani, K. Esmaeilpour, M. Nazeri, S. Mazhari and V. Sheibani, 2015. Exercise improves learning and memory impairments in sleep deprived female rats. Physiol. Behav., 138: 285-291.
08. Allison, T. and D.V. Cicchetti, 1976. Sleep in mammals: Ecological and constitutional correlates. Science, 194: 732-734.
09. Zou, H., 2006. The adaptive lasso and its oracle properties. J. Am. Stat. Assoc., 101: 1418-1429.