A Comparison between Linear Regression – Lasso Quantile Regression Methods in Selecting Best Subset Variables

Neveen Sayed- Ahmed

Assistant Lecturer of Statistics, Faculty of Commerce, Al-Azhar University (Girls' Branch)

Dr. Elham Abdul-Razik Ismail

Associate Professor of Statistics, Faculty of Commerce, Al-Azhar University (Girls' Branch)

Keywords: Quantile Regression – Linear Lasso – Selection of Variables – backward regression – forward regression – Kruskal-Wallis test.

ABSTRACT

One of the main topics in the development of predictive models is the identification of variables which are predictors of a given outcome. Automated model selection methods, such as backward or forward stepwise regression, are classical solutions to this problem, but are generally based on strong assumptions about the functional form of the model or the distribution of residuals. The quantile regression can give complete information about the relationship between the response variable and covariates on the entire conditional distribution, and has no distributional assumption about the error term in the model. The study aimed to: 1- evaluate the performance of the Lasso regression as a good alternative to ordinary least squares (OLS) and least absolute value (LAV) regression methods when used to estimate the regression coefficients. 2- Demonstrate the efficiency of the Lasso regression when used to select the best subset variables. 3- present a numerical application to demonstrate the efficiency of the Lasso quantile regression when different quantile regression values are used to select the best subset of variables and estimation regression coefficients. The study results showed that Lasso regression is an appropriate model for estimating the parameters and selection of variables. Lasso quantile regression as regularization technique for simultaneous estimation and variable selection methods are often highly time consuming and maybe suffer from instability.

1- INTRODUCTION

Variable selection plays an important role in classification. Before beginning designing a classification method, when many variables are involved, only those variables that are really required should be selected; that is, the first step is to eliminate the less significant variables from the analysis. There can be many reasons for selecting only a subset of the variables instead of the whole set of candidate variables: (1) It is cheaper to measure only a reduced set of variables, (2) Prediction accuracy may be improved through exclusion of redundant and irrelevant variables, (3) The predictor to be built is usually simpler and potentially faster when fewer input variables are used and (4) Knowing which variables are relevant can give insight into the nature of the prediction problem and allows a better understanding of the final classification model. Research in variable selection started in the early 1960s. Over the past four decades, extensive research into feature selection has been conducted. Much of the work is related to medicine and biology. The selection of the best subset of variables for building the predictor is not a trivial question, because the number of subsets to be considered grows exponentially with the number of candidate variables. Even with a moderate number of candidate variables, not all the possible subsets can be evaluated, which means that feature selection is a NP (Nondeterministic Polynomial) -Hard computational problem. This means that when the size of the problem is large finding an optimum solution in practice is not feasible (Casado et. al 2007).

The ordinary least squares (OLS) method is one of the oldest and most widely used statistical tools for linear models. Its theoretical properties have been extensively studied and are fully understood. Despite its many superior properties, the LS estimate can be sensitive to outliers and, therefore, non-robust. Its performance in terms of accuracy and statistical inferences may be compromised when the errors are large and heterogeneous. The least absolute deviation (LAD) method, which is also known as the L1 method and has an equally long history, provides a useful and plausible alternative (Birkes and Dodge1993).

The least absolute deviation (LAD) method is a widely known alternative to the classical least squares (OLS) method for statistical analysis of linear regression models. Instead of minimizing the sum of squared errors, it minimizes the sum of absolute values of errors (Arthanari and Dodge 1993).

Quantile regression is a statistical technique intended to estimate, and conduct inference about the conditional quantile functions. Just as the classical linear regression methods estimate models for conditional mean function, quantile regression offers a mechanism for estimating models for conditional median function, and the full range of other conditional quantile functions (Wu and Liu 2009).

In a previous study to evaluate the performance of two linear lasso (L_1-Lasso, L_2-Lasso) methods. The two methods are used to select the best subset of variables and estimate the parameters of the quantile regression equation when four error distributions, with two different sample sizes and two different parameters for each error distribution. All results showed that the L_1-Lasso, and L_2-Lasso linear methods are the same and much better with fat-long tailed distribution. (Gharib2013).

The study aimed to evaluate the performance of the Lasso regression as a good alternative to Ordinary least squares (OLS) and least absolute deviation (LAD) regression methods when used to estimate the regression coefficients. Demonstrate the efficiency of the Lasso regression when used to select the best subset variables. Present a numerical application to demonstrate the efficiency of the Lasso quantile regression as regularization technique for simultaneous estimation and variable selection when different quantile values are used.

2- LASSO

This section is concerned with the procedures which are used in the present work Lasso and Lasso quantile regression. A method that has received a great deal of attention in the statistics literature is the least absolute shrinkage selection operator (LASSO) of Tibshirani (1996). The main difference between a LASSO and a ridge regression is the use of a L1 instead of an L2 penalty. This difference turns out to important because an L2 penalty only shrinks coefficients to zero but never sets them to zero exactly.

Since Tibshirani (1996) proposed the least absolute shrinkage and selection operator lasso, which can effectively select important explanatory variables and estimate regression parameters simultaneously. The combination of the quantile regression and Lasso penalty is computationally easy to implement via the standard linear programming. Simulation studies are conducted to assess the finite sample performance of the proposed method. In the general linear model with independent and identically distributed errors, the least absolute deviation (LAD) or L_1 method has been a viable alternative to the least squares method especially for its superior robustness properties. Consider the linear regression model,

$$Y_i = \beta' x_i + e_i \qquad 1 \le i \le n$$

where x_i are known p-vectors, β the unknown p-vector of regression coefficients, and e_i the i.i.d random errors. The L_1 estimator $\hat{\beta}_{L_1}$ is defined as a minimize of the L_1 loss function

$$L_n(\beta) = \sum_{i=1}^{n} \left| Y_i - \beta' x_i \right|$$

As Eforn et al.(2004) the least squares estimate $\beta^{Ls} = (X'X)^{-1} X'Y$ uniquely minimizes the squared loss

$$\sum_{i=1}^{n} (Y_i - \beta' x_i)^2$$

Lasso estimate is defined as the minimum of

where $0 \le s \le 1$ controls the amount of shrinkage that is applied to the estimates.

Tibshirani (1996) proposed the least absolute shrinkage and selection operator lasso as follows:

$$\sum_{i=1}^{n} \left| Y_i - \beta' x_i \right| + \lambda \sum_{j=1}^{n} \| \beta_j \|_1$$

and

$$\sum_{i=1}^{n} |Y_i - \beta' x_i| + \lambda \sum_{j=1}^{n} \| \beta_j \|_2$$

where $\| \beta_j \|_1$ is the usual L_1 estimator and $\| \beta_j \|_2$ is the usual L_2 estimator. The tuning parameter β there plays a crucial role of striking a balance between estimation of β_j and variable selection, λ is the tuning parameter.

The study, a parallel approach borrowing is proposed the ideas from Lasso by using the L_1 penalty and L_2 penalty, but with the least squares loss replaced by the L_1 loss in quantile regression model. In doing so, we gain advantages in two fronts. First, it allows us to penetrate the difficult problem of variable selection for the L_1 regression. Appealingly, the shrinkage property of the Lasso estimator continues to hold in L_1 regression. Second, the single criterion function with both components

being of L_1-type reduces (numerically) the minimization to a strictly linear programming problem, making any resulting methodology extremely easy to implement. To be specific, our proposed estimator is a minimize of the following criterion function

$$\sum_{i=1}^{n} \rho |Y_i - \beta' x_i| + \lambda \sum_{j=1}^{n} \| \beta_j \|_1$$

and

$$\sum_{i=1}^{n} \rho |Y_i - \beta' x_i| + \lambda \sum_{j=1}^{n} \| \beta_j \|_2$$

where $\| \beta_j \|_1$ is the usual L_1 estimator and $\| \beta_j \|_2$ is the usual L_2 estimator. The tuning parameter β there plays a crucial role of striking a balance between estimation of β_j and variable selection, λ is the tuning parameter, and ρ is quantile values (Koenker2004).

3- VARIABLE SELECTION PROCEDURES

This section is concerned with the traditional variable selection methods which are used in the present work. Three popular iterative search algorithms for choosing a "best subset" regression are **forward selection**, **backward elimination**, and **stepwise regression**. In contrast to all subset searches based on a goodness-of-fit criterion, these algorithms are called "directed search" algorithms because they avoid all subset searches by following certain rules in conducting the search. The study briefly summarizes these methods.

Forward selection method is often used to provide an initial screening of the candidate variables when a large group of variables exists. For example, suppose you have fifty to one hundred variables to choose from, way outside the realm of the all-possible regressions procedure. A reasonable approach would be to use this forward selection procedure to obtain the best ten to fifteen variables and then apply the all-possible algorithm to the variables in this subset. This procedure is also a good choice when multicollinearity is a problem. The forward selection method is simple to define. You begin with no candidate variables in the model. Select the variable that has the highest R-Squared. At each step, select the candidate variable that increases R-Squared the most. Stop adding variables when none of the remaining variables are significant. Note that once a variable enters the model, it cannot be deleted.

Backward selection method is less popular because it begins with a model in which all candidate variables have been included. However, because it works its way down instead of up, you are always retaining a large value of R-Squared. The problem is that the models selected by this procedure may include variables that are not really necessary. The user sets the significance level at which variables can enter the model. The backward selection model starts with all candidate variables in the model. At each step, the variable that is the least significant is removed. This process continues until no non-significant variables remain. The user sets the significance level at which variables can be removed from the model (Xu and Zhang2001).

3-NUMERICAL APPLICATION

This section discusses the numerical application which used to evaluate the performance of the Lasso regression method when used to estimate and select the best subset variables. Ordinary least squares (OLS) and least absolute deviation (LAD) methods give nonzero estimates to all coefficients. The Lasso is a regression method similar to ordinary least squares (OLS) and least absolute deviation (LAD) regression methods. Lasso minimizes the Residual Sum of Squares (*RSS*) but poses a constraint to the sum of the absolute values of the coefficients being less than a constant. This additional constraint is moreover similar to that introduced in Ridge regression, where the constraint is to the sum of the squared values of the coefficients. This simple modification allows Lasso to perform also variable selection because the shrinkage of the coefficients is such that some coefficients can be shrunk exactly to zero. The aimed of this study are:

1- Establishing the Lasso regression is a good alternative to Ordinary least squares (OLS) and least absolute deviation (LAD) regression methods.
2- Justifying the Lasso regression is a good alternative to traditionally variable selection methods (forward and backward method).
3- Realizing the importance of employing quantile regression method in the analysis of real data.

This section presents a numerical application to evaluate the performance of the Lasso and Lasso Quantile regression methods. The evaluation has been done by three steps:

1- Comparing between Ordinary least squares (OLS), least absolute deviation (LAD) and Lasso regression methods when the three methods are used to estimate the parameters in regression model.
2- Comparing between Lasso regression, forward and backward methods when the three methods are used in selection of best subset of variables in regression model.
3- Comparing between Lasso regression and Lasso quantile regression parameters when different quantile regression values are used to select the best

subset of variables and estimation the regression coefficients.

The study uses the data set reported in Jobson 1991. The data was collected from the DATASTREAM database for a sample of 40 UK listed companies. The observations obtained from 13 financial variables. This data is used to estimate a linear relation between the return on capital employed (RETCAP) and remaining 12 variables. Table (1) listed the 13 variables which used in the study. This study introduced a program by using GAMS 2.25 statistical package to calculate Ordinary least squares (OLS), least absolute deviation (LAD), Lasso and Lasso quantile regression estimators.

4- RESULTS

First: comparing between Ordinary least squares (OLS), least absolute value deviation (LAD) and Lasso regression methods

This section aims to discuss the results of the comparison between ordinary least squares (OLS), least absolute value deviation (LAD) and Lasso regression methods when the three methods are used to estimate regression coefficients. Table (2) presents the estimated regression coefficients for UK financial accounting data.

ordinary least squares (OLS) and least absolute deviation (LAD) methods give nonzero estimates to all coefficients. The Lasso gave zero coefficients to capint, logsale, invtast and fattot; subset selection and gave non-zero coefficients to the rest. The study tested the statistical significant with Kruskal-Wallis rank – sum test. Kruskal-Wallis statistical test is used to compare between three calculated coefficients for OLS, LAD and Lasso methods when the three methods are used to estimate the regression coefficients. The data provides statistically significant (p-value=0.8972> 0.05) for three calculated coefficients or the Lasso is a regression method similar to Ordinary Least Squares (OLS) and least absolute deviation (LAD) regression methods when it is used to estimate regression coefficients. That means Lasso regression estimators is a good alternative to Ordinary least squares (OLS), least absolute deviation (LAD) estimators.

Table (2) estimated regression coefficients for UK financial accounting data

variables	OLS coefficients	LAD coefficients	Lasso coefficients
GREARRAT	-0.027	0.049	-0.87
CAPINT	-4.26E-4	-0.002	0
WCFTDT	0.478	0.379	0.859
LOGSALE	0.101	0.095	0
LOGASST	-0.028	-0.036	-1.76
CURRAT	-0.214	-0.175	.273
QUIKRAT	0.164	0.136	-0.116
NFATAST	-0.360	-0.372	0.092
INVTAST	0.273	0.225	0
FATTOT	-0.089	-0.034	0
PAYOUT	-0.015	-0.014	0.342
WCFTCL	0.069	0.100	0.144

Second: comparing between Lasso regression, forward and backward methods

This section is to discuss the results of the comparison between Lasso regressions, forward and backward methods when the three methods are used to select the best subset of variables in regression model. As an illustration for UK financial accounting data which reported in Jobson 1991. The results of the backward process are similar to the results from the forward procedure except that the procedure is reversed. The order of entry in the forward selection method was wcftdt, quirat, nfatast, logsale, currt, logasst, wcftct, fattot, payout and invtast. The order of entry in the forward selection method was gearrat, capint, invtast, wcftdt, payout, fattot, logasst and quikrat. Table (3) presents the best subset or non – zero variables (estimated regression coefficients) for UK financial accounting data by three methods forward, backward and Lasso methods.

Table (3) estimated regression coefficients for UK financial accounting data

Lasso coefficients	Backward coefficients	forward coefficients	variables
-0.87	0	0	GREARRAT
0	0	0	CAPINT
0.859	0.351	0.361	WCFTDT
0	0.100	0.100	LOGSALE
-1.76	-0.061	-0.062	LOGASST
.273	-0.176	-0.176	CURRAT
-0.116	0.124	0.124	QUIKRAT
0.092	-0.349	-0.350	NFATAST
0	0.161	0.161	INVTAST
0	-0.109	-0.109	FATTOT
0.342	-0.016	-0.016	PAYOUT
0.144	0.198	0.199	WCFTCL
0.777	0.780	0.780	R^2- VALUE

The results of the backward process are similar to the results from the forward procedure except that the procedure is reversed. So the study used Wilcoxon statistical test is used to compare between two calculate coefficients for forward or backward and Lasso methods when the three methods are used to select the best subset variables. The data provides statistically significant (p-value=0.9641> 0.05) for three calculated coefficients or the Lasso is a regression method similar to forward, backward elimination regression methods when it is used to select the best subset variables. That means Lasso regression estimators are a good alternative to traditionally variable selection methods (forward and backward method).

Third: comparison between Lasso regression and Lasso quantile regression parameters when different quantile regression values

This section is to discuss the results of the comparison between Lasso regression and Lasso quantile regression parameters when different quantile regression values are used in selection of best subset of variables and estimation the parameters for regression model. The median is a special quantile, one that describes the central location of a distribution. Conditional-median regression is a special case of quantile regression in which the conditional .5th quantile is modeled as a function of covariates. More generally, other quantiles can be used to describe noncentral positions of a distribution. The *quantile* notion generalizes specific terms like quartile, quintile, decile, and percentile. The study suggested the following values of quantiles are $\rho = (0.1; 0.25; 0.8)$. Table (4) presents the estimated regression coefficients, best subset for UK financial accounting data when three different quantile values are used. Kruskal-Wallis statistical test is used to compare between four calculate coefficients for Lasso and Lasso quantile methods when the two methods are used to estimate the regression coefficients and select the best subset variables. The Lasso quantile method applied with three different quantile values. The data provides statistically significant (p-value=0.9120> 0.05) for four calculated coefficients or the Lasso quantile regression is a natural extension of the linear regression model, estimate, select best subset variables and exibility in assessing the effect of predictors on different locations of the response distribution. Lasso quantile regression offers a mechanism for estimating models for conditional median function, and the full range of other conditional quantile functions.

Table (4) estimated regression coefficients for UK financial accounting data when three different quantile values are used.

Lasso coefficients	Q Lasso $(\rho = 0.8)$ coefficients	Q Lasso $(\rho = 0.25)$ coefficients	Q Lasso $(\rho = 0.1)$ coefficients	variables
-0.87	0.319	0.103	0.258	GREARRAT
0	-0.578	-0.578	-0.635	CAPINT
0.859	0.041	0.041	-0.066	WCFTDT
0	0.007	0.063	0.270	LOGSALE
-1.76	-0.022	-0.037	-0.131	LOGASST
.273	0	0.351	0.055	CURRAT
-0.116	0.005	-0.342	0	QUIKRAT
0.092	0	0	0	NFATAST
0	0	0	0	INVTAST
0	0	0	-0.701	FATTOT
0.342	-0.023	-0.023	0.051	PAYOUT
0.144	0	0	0	WCFTCL

5- Conclusion

1- Lasso and Lasso quantile regression method are considered modify and improve methods for the traditional statistical methods which are used to estimate the parameter of the linear regression models and the statistical method which used to selection variables.

2- This study aims to introduce the quantile regression model to a broad audience of social scientists who are interested in modeling both the location and shape of the distribution they wish to study. It is also for researchers who are concerned about the sensitivity of linear regression models to skewed distributions and outliers.

3- Lasso quantile regression as regularization technique for simultaneous estimation and variable selection when different quantile values are used.

4- The study results showed that Lasso regression is an appropriate model for estimating the parameters and selection of variables.

5- Lasso quantile regression as regularization technique for simultaneous estimation and variable selection methods are often highly time consuming and maybe suffer from instability.

6- Quantile Regression is a good alternative to ordinary least squares regression. Whereas the sum of squared errors is minimized in ordinary least squares regression, the median regression estimator minimized the sum of absolute errors.

7- The lasso idea is quite general and can be applied in a variety of statistical models: extensions to generalized regression models and tree-based models are briefly described.

Table (1) financial accounting data for 40 UK companies

CURRT	LOGASST	LOGSALE	WCFTCT	CAPINT	GEARRAT	RETCAP
1.53	4.3	4.11	0.25	0.64	0.46	0.26
1.73	4	4.25	0.33	1.79	0	0.57
0.44	4.88	4.44	0.2	0.36	0.24	0.09
1.23	4.44	4.71	0.21	1.86	0.45	0.32
1.76	4.75	4.85	0.12	1.26	0.91	0.17
1.44	5.42	5.61	0.25	1.54	0.26	0.24
0.83	4.3	4.83	0.4	3.34	0.52	0.53
1.45	4.35	4.49	0.37	1.38	0.24	0.26
2.89	4.17	4.13	0.21	0.91	0.19	0.13
2.13	4.17	4.4	0.18	1.7	0.29	0.16
1.1	4.09	4.3	0.01	1.6	0.85	0.06
4.57	4.45	3.62	0.7	0.15	0.02	0.07
0.47	4.35	4.13	-0.32	0.6	0.76	-0.18
0.85	3.74	4.11	0.11	2.34	0.39	0.12
1.81	4.55	4.63	0.65	1.19	0.06	0.15
12.98	4.18	0	1.47	0	0	0.03
1.43	4.01	4.06	0.08	1.12	0.39	0.08
1.75	4.06	4.21	0.13	1.42	0.26	0.09
1.49	3.62	3.99	0.23	2.33	0.15	0.25
1.35	4.3	4.51	-0.07	1.62	0.67	-0.03
0.29	4.24	1.74	0.04	0	0.15	0.03
1.42	4.02	4.24	0.03	1.65	0.34	0.04
1.12	3.41	3.52	0.14	1.29	0.38	0.17
1.5	3.96	4.03	0.05	1.2	0.18	0.07
1.3	3.97	4.35	0.09	2.4	0.45	0.11
1.08	4.1	4.72	-0.01	3.46	0.54	0.04
1.32	3.93	4.26	0.46	2.11	0.09	0.04
1.43	4.6	4.67	0.17	1.16	0.17	0.11
1.32	4.38	4.82	0.08	2.75	0.35	0.14
1.71	4.86	5.1	0.34	1.88	0.13	0.29
1.97	3.96	4.14	0.15	1.53	0.13	0.02
0.91	4.53	5.6	0.12	3.42	0.59	0.1
1.09	6.25	6.29	0.26	1.11	0.16	0.14
1.45	4.46	4.77	0.23	2.03	0.17	0.11
1.83	6.18	6.23	0.26	1.1	0.43	0.29
1.23	4.47	4.3	0.05	0.66	0.37	0.4
2.35	4.05	4.27	0.23	1.68	0.04	0.17
1.88	5.01	5.24	0.25	1.71	0.17	0.16
1.37	3.85	3.99	0.39	1.8	0.02	0.14
2.49	4.37	4.54	0.2	1.49	0.23	0.13

Table (1) financial accounting data for 40 UK companies (continued)

WCFTCL	PAYOUT	FATTOT	INVTAST	NFATAST	QUIKRAT
0.25	0.07	0.12	0.74	0.1	0.18
0.33	0.3	0.15	0.27	0.12	1.26
0.5	0.57	0.97	0.01	0.94	0.39
0.23	0	0.52	0.29	0.29	0.69
0.21	0.31	0.54	0.33	0.26	0.9
0.37	0.15	0.57	0.06	0.42	1.23
0.59	0.21	0.21	0	0.14	0.83
0.44	0.16	1.04	0.36	0.4	0.58
0.21	0.39	0.11	0.29	0.06	1.95
0.21	0.46	0.4	0.58	0.21	0.56
0.01	0	0.38	0.34	0.23	0.73
0.7	0	0.63	0	0.54	4.51
-0.58	0	0.84	0	0.54	0.47
0.11	0	0.97	0.49	0.41	0.14
0.81	0.26	0.77	0.1	0.65	1.25
1.47	0	0.06	0	0.05	12.98
0.09	0	0.44	0.36	0.36	0.59
0.13	0.6	0.41	0.33	0.31	0.92
0.23	0.23	0.49	0.37	0.21	0.79
-0.07	0	0.41	0.43	0.2	0.57
0.04	0	0	0	0	0.29
0.3	4.21	0.43	0.36	0.28	0.71
0.15	0.16	0.37	0.33	0.27	0.61
0.05	1.66	0.31	0.31	0.21	0.9
0.1	0.35	0.6	0.22	0.36	0.86
-0.2	0	0.51	0.19	0.49	0.66
0.55	0.97	1.16	0.15	0.64	0.78
0.17	0.71	0.55	0.23	0.32	0.76
0.09	0.56	0.25	0.61	0.21	0.3
0.37	0.43	0.5	0.3	0.29	0.96
0.19	0	0.85	0.18	0.43	1.33
0.14	0.65	0.57	0.2	0.41	0.58
0.39	0.47	0.81	0.11	0.67	0.59
0.26	0.67	0.84	0.27	0.45	0.75
0.42	0.52	0.43	0.2	0.28	1.17
0.05	1.83	0.53	0.39	0.47	0.33
0.24	0.71	0.26	0.39	0.11	1.33
0.32	0.31	0.65	0.31	0.29	1.05
0.39	0.78	0.82	0.24	0.52	0.66
0.25	0.58	0.46	0.28	0.32	1.47

6-References

1- Arthanari, T. S. and Dodge, Y. (1993), *Mathematical programming in statistics*. John Wiley and Sons, Inc. New York.
2- Birkes, D. and Dodge, Y. (1993), *Alternative methods of regression*. John Wiley and Sons, Inc. New York.
3- Casado S., Pacheco J. y Núñez L. (2007), *A new variable selection method for classification.*JEL (Journal Economic Literature) Classification: C44, C61.
4- Efron, B., Hastie, T., Johnstone, I. and Tibshirani, R. (2004), *Least angle regression.* The Annals of Statistics, 32, 407-499.
5- Gharib, A. N. (2013), *Linear Programming Approach for Selection of Variables in Quantile. Regression* Unpublished Ph.D. thesis, Faculty of Commerce, Al-Azhar University-Girls' Branch.
6- Jobson, J.D. (1991), *Applied Multivariate Data Analysis*. Vol 1. Springer Verlag, Inc. New York.
7- Koenker, R. (2004), *Quantile regression for longitudinal data*. *Journal of* Multivariate Analysis, 91, 74-89.
8- Tibshirani, R. (1996), *Regression shrinkage and selection via the Lasso*. Journal of the Royal Statistical Society, Series B, 58, 267-288.
9- Wu, Y. and Liu, Y. (2009), *Variable selection in quantile regression.* Statistica Sinica 19, 801-817.
10- Xu, L. and Zhang, H. (2001), *Comparison of different methods for variable selection*. Analytica Chimica Acta, 446, 477-483.