Group Identification and Variable Selection in Quantile Regression

Ali Alkenani and Basim Shlaibah Msallam

1Department of Statistics, College of Administration and Economics, University of Al-Qadisiyah, Iraq
2Supervision & Scientific Evaluation Office, Ministry of Higher Education and Scientific Research, Iraq

Correspondence should be addressed to Ali Alkenani; ali.alkenani@qu.edu.iq

Received 24 November 2018; Revised 21 February 2019; Accepted 28 March 2019; Published 10 April 2019

Using the Pairwise Absolute Clustering and Sparsity (PACS) penalty, we proposed the regularized quantile regression QR method (QR-PACS). The PACS penalty achieves the elimination of insignificant predictors and the combination of predictors with indistinguishable coefficients (IC), which are the two issues raised in the searching for the true model. QR-PACS extends PACS from mean regression settings to QR settings. The paper shows that QR-PACS can yield promising predictive precision as well as identifying related groups in both simulation and real data.

1. Introduction

The regression model is one of the most important statistical models. The ordinary least squares regression (OLS) estimates the conditional mean function of the response. The least absolute deviation regression (LADR) estimates the conditional median function and it is resistant to outliers. The QR was introduced by Koenker and Bassett [1] as a generalization of LADR to estimate the conditional quantile function of the response. Consequently, QR gives us much more information about the conditional distribution of the response. QR has attracted a vast amount of interest in literature. It is applied in many different areas such as economics, finance, survival analysis, and growth chart.

Variable selection (VS) is very important in the process of model building. In many applications, the number of variables is huge. However, keeping irrelevant variables in the model is undesirable because it makes the model difficult to interpret and may affect negatively its ability of prediction. Many different penalties were suggested to achieve VS. For example, Lasso [2], SCAD [3], fused Lasso [4], elastic-net [5], group Lasso [6], adaptive Lasso [7], adaptive elastic-net [8], and MCP [9].

Under QR framework, Koenker [10] combined the Lasso with the mixed-effect QR model to encourage shrinkage in estimating the random effects. Wang, Li, and Jiang [11] combined LADR with the adaptive Lasso penalty. Li and Zhu [12] proposed L1-norm penalized QR (PQR) by combining QR with Lasso penalty. Wu and Liu [13] introduced PQR with the SCAD and the adaptive Lasso penalties. Slawski [14] proposed the structured elastic-net regularizer for QR.

In the setting \( p > n \), where \( p \) represents the number of predictors and \( n \) represents the sample size, Belloni and Chernozhukov [15] studied the theory of PQR for the Lasso penalty. They considered QR in high-dimensional sparse models. Wang, Wu, and Li [16] investigated the methodology of PQR in ultrahigh dimension for the nonconvex penalties such as SCAD or MCP. Peng and Wang [17] proposed and studied a new iterative coordinate descent algorithm for solving nonconvex PQR in high dimension.

The search for the true model focuses on two issues: deleting irrelevant predictors and merging predictors with IC [18]. Although the above penalties can achieve the first issue, they fail in achieving the second one. The two issues can be achieved through Pairwise Absolute Clustering and
Sparsity (PACS) [18]. Moreover, PACS is an oracle method for simultaneous group identification and VS.

The limitations of existing variable selection methods motivate the authors to write this paper. The aim of the current research is to find an effective procedure for simultaneous group identification and VS under QR framework.

In this paper, we suggested the QR-PACS to get the advantages over the existing PQR methods. QR-PACS benefits from the ability of PACS on achieving the mentioned issues of the discovery of the true model which is unavailable in Lasso, Adaptive Lasso, SCAD, MCP, Elastic-net, and structured elastic-net.

The rest of the paper is organized as follows. In Section 2, penalized linear QR is reviewed briefly. QR-PACS is introduced in Section 3. The numerical results of simulations and real data are presented in Sections 4 and 5, respectively. The conclusions are reported in Section 6.

2. Penalized Linear QR

QR is a widespread technique used to describe the distribution of an outcome variable ($y_i$), given a set of predictors ($x_i$). Let $x_i$ be a $p \times 1$ vector of predictors for the $i$th observation and $q_r(x_i)$ be the inverse cumulative distribution function of $y_i$ given $x_i$. Then, $q_r(x_i) = x_i^T \beta_r = \sum_{j=1}^{p} x_{ij} \beta_{jr}$, where $\beta_r$ is a vector of $p$ unknown parameters and $r$ is the level of quantile.

Koenker and Bassett [1] suggested estimating $\tilde{\beta}_r$ as follows:

$$\min_{\beta_r} \sum_{i=1}^{n} \rho_r \left( y_i - \sum_{j=1}^{p} x_{ij} \beta_{jr} \right),$$

where $\rho_r(\cdot)$ is the check loss function defined as

$$\rho_r(u) = \tau u I_{[0,\infty)}(u) - (1 - \tau) u I_{(-\infty,0)}(u).$$

Under regularization framework, Li and Zhu [12], Wu and Liu [13], Slawski [14], and Wang, Wu, and Li [16] among others proposed the penalized versions of (1) by adding different penalties as follows:

$$\min_{\beta_r} \sum_{i=1}^{n} \rho_r \left( y_i - \sum_{j=1}^{p} x_{ij} \beta_{jr} \right) + \lambda \sum_{j=1}^{p} P(\beta_{jr})$$

where $\lambda > 0$ is the penalization parameter and $P(\beta_{jr})$ is the penalty function.

For the rest of this paper, the subscript $\tau$ is omitted for notational convenience.

3. Penalized Linear QR through PACS (QR-PACS)

In this section, we incorporate PACS into the optimization of (1) to propose QR-PACS. Under the QR setup, the predictors $x_{ij}$ are standardized and the response $y_i$ is centered, $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, p$. The QR-PACS is proposed for simultaneous group identification and VS in QR. The QR-PACS encourages correlated variable to have equal coefficient values. The equality of coefficients is attained by adding group identification penalty to the pairwise differences and sums of coefficients. The QR-PACS estimates are proposed as minimizers of

$$\min_{\beta_r} \sum_{i=1}^{n} \rho \left( y_i - \sum_{j=1}^{p} x_{ij} \beta_{jr} \right) + \lambda \sum_{j=1}^{p} \omega_{jr} |\beta_{jr}|$$

$$+ \sum_{1 \leq j < k \leq p} \omega_{jk} (|\beta_k - \beta_j| + \sum_{1 \leq j < k \leq p} \omega_{jk} (|\beta_k + \beta_j|),$$

where $\omega$ are the nonnegative weights.

The PACS penalty in (4) consists of $\lambda \sum_{j=1}^{p} \omega_{jr} |\beta_{jr}|$ that encourages sparseness, $\lambda \sum_{1 \leq j < k \leq p} \omega_{jk} (|\beta_k - \beta_j| + \sum_{1 \leq j < k \leq p} \omega_{jk} (|\beta_k + \beta_j|)$, which are employed for the group identification and encourage equality of coefficients. The second term of the penalty encourages the same sign coefficients to be set as equal, while the third term encourages opposite sign coefficients to be set as equal in magnitude.

Choosing appropriate adaptive weights is very important for PACS. In QR-PACS, we employed the adaptive weights that incorporate correlations into the weights as suggested by Sharma et al. [18] with a small modification as follows:

$$\omega_j = \left[ \tilde{\beta}_j \right]^{-1},$$

$$\omega_{jk} = \left( 1 - r_{jk} \right)^{-1} \left[ \tilde{\beta}_k - \tilde{\beta}_j \right]^{-1}$$

$$\text{and } \omega_{jk} = \left( 1 + r_{jk} \right)^{-1} \left[ \tilde{\beta}_k + \tilde{\beta}_j \right]^{-1}$$

for $1 \leq j < k \leq p$.

where $\tilde{\beta}$ is a $\sqrt{n}$ consistent estimator of $\beta$, such as the PACS [18] estimates or other shrinkage QR estimates, and $r_{jk}$ is the biweight midcorrelation ($j,k$)th pair of predictors. We propose to employ the biweight midcorrelation [19, 20] instead of Pearson correlation which is used in the adaptive weights in [18] to obtain robust correlation and robust weights.

In this paper, ridge quantile estimates were employed as initial estimates for $\beta$’s to obtain weights performing well in studies with collinear predictors.

4. Simulation Study

In this section, five examples were carried out to assess QR-PACS method by comparing it with existing selection approaches under QR setting in both prediction precision and model discovery. A regression model was generated as follows.

$$y = X\beta + \sigma \epsilon$$

In all examples, predictors $X$ and the error term $\epsilon$ were standard normal.
Table 1: ME, SA, GA, and SGA results of Example 1 for n=100.

| r   | Criterion | Ridge QR | QR-LASSO | QR-SCAD | QR-adaptive LASSO | QR-elastic-net | QR-PACS |
|-----|-----------|----------|----------|---------|--------------------|----------------|---------|
| 0.10| ME(S.E)   | 0.1144(0.0176) | 0.0834(0.0182) | 0.0684(0.0165) | 0.0673(0.0153) | 0.0623(0.0112) | 0.0432(0.0093) |
|     | SA        | 0        | 62       | 79       | 80                 | 87             | 79      |
|     | GA        | 0        | 0        | 0        | 0                  | 0              | 88      |
|     | SGA       | 0        | 0        | 0        | 0                  | 0              | 72      |
| σ=1 | ME(S.E)   | 0.0832(0.0056) | 0.0524(0.0065) | 0.0443(0.0056) | 0.0431(0.0050) | 0.0404(0.0041) | 0.0143(0.0020) |
|     | SA        | 0        | 64       | 81       | 82                 | 86             | 80      |
|     | GA        | 0        | 0        | 0        | 0                  | 0              | 89      |
|     | SGA       | 0        | 0        | 0        | 0                  | 0              | 73      |
| 0.50 | ME(S.E)  | 0.1140(0.0173) | 0.0829(0.0177) | 0.0672(0.0150) | 0.0668(0.0146) | 0.0619(0.0105) | 0.0429(0.0082) |
|     | SA        | 0        | 63       | 81       | 81                 | 88             | 80      |
|     | GA        | 0        | 0        | 0        | 0                  | 0              | 89      |
|     | SGA       | 0        | 0        | 0        | 0                  | 0              | 72      |
| 0.75 | ME(S.E)  | 0.1471(0.0183) | 0.1168(0.0188) | 0.1062(0.0162) | 0.1056(0.0162) | 0.0953(0.0120) | 0.0755(0.0106) |
|     | SA        | 0        | 59       | 77       | 77                 | 84             | 77      |
|     | GA        | 0        | 0        | 0        | 0                  | 0              | 86      |
|     | SGA       | 0        | 0        | 0        | 0                  | 0              | 70      |
| σ=3 | ME(S.E)   | 0.1153(0.0066) | 0.0855(0.0071) | 0.0763(0.0060) | 0.0756(0.0060) | 0.0730(0.0049) | 0.0473(0.0034) |
|     | SA        | 0        | 61       | 80       | 80                 | 83             | 78      |
|     | GA        | 0        | 0        | 0        | 0                  | 0              | 87      |
|     | SGA       | 0        | 0        | 0        | 0                  | 0              | 71      |
| 0.50 | ME(S.E)  | 0.1466(0.0179) | 0.1161(0.0185) | 0.1056(0.0161) | 0.1044(0.0158) | 0.0945(0.0117) | 0.0744(0.0099) |
|     | SA        | 0        | 61       | 79       | 79                 | 85             | 78      |
|     | GA        | 0        | 0        | 0        | 0                  | 0              | 87      |
|     | SGA       | 0        | 0        | 0        | 0                  | 0              | 70      |

We compared QR-PACS with ridge QR, QR-Lasso, QR-SCAD, QR-adaptive Lasso, and QR-elastic-net. The performance of the methods was compared using model error (ME) criterion for prediction accuracy which was defined by \((\hat{\beta} - \beta)^\top V(\hat{\beta} - \beta)\) where \(V\) represents the population covariance matrix of \(X\) and the resulting model complexity for model discovery. The median and standard error (SE) of ME were reported. Also, selection accuracy (SA, % of true models identified), grouping accuracy (GA, % of true groups identified), and % of both selection and grouping accuracy (SGA) were computed and reported. Note that none of ridge QR, QR-Lasso, QR-SCAD, QR-adaptive Lasso, and QR-elastic-net perform grouping. The sample size was 100 and the simulated model was replicated 100 times. Some typical examples are reported as follows.

Example 1. In this example, we assumed the true parameters for the model of study as \(\beta = (2,2,2,0,0,0,0,0)^\top, X \in \mathbb{R}^8\). The first three predictors were highly correlated with correlation equal to 0.7 and their coefficients were equal in magnitude, while the rest were uncorrelated.

Example 2. In this example, the true coefficients were assumed as \(\beta = (0.5,1,2,0,0,0,0,0)^\top, X \in \mathbb{R}^8\). The first three predictors were highly correlated with correlation equal to 0.7 and their coefficients were equal in magnitude, while the rest were uncorrelated.

Example 3. In this example, the true parameters were \(\beta = (1,1,0.5,1,2,0,0,0,0)^\top, X \in \mathbb{R}^{10}\). The first three predictors were highly correlated with correlation equal to 0.7 and their coefficients were equal in magnitude, while the correlation for the second three predictors was equal to 0.3 and their coefficients were different in magnitudes. The remaining predictors were uncorrelated.

Example 4. In this example, the true parameters were assumed as \(\beta = (1,1,1,0.5,1,2,0,0,0,0)^\top, X \in \mathbb{R}^{10}\). The first three predictors were correlated with correlation equal to 0.3 and their coefficients were equal in magnitude, while the correlation for the second three predictors was 0.7 and their coefficients were different in magnitudes. The remaining predictors were uncorrelated.

Example 5. In this example, the true parameters were assumed as \(\beta = (2,2,2,1,0,0,0,0,0,0)^\top, X \in \mathbb{R}^{10}\). The first three and the second two predictors were highly correlated with correlation equal to 0.7 and their coefficients were different in magnitude, while the rest were uncorrelated.

For all the values of \(r\) and \(\sigma\), Table 1 shows that the QR-PACS method has the lowest ME. Although the QR-elastic-net and QR-adaptive Lasso have the highest SA, it is clear that all the considered methods do not perform grouping except QR-PACS. The QR-PACS method successfully identifies the groups of predictors as seen in the GA and SGA rows.
In Table 2, the percentage of no-grouping (NG, no groups found) and percentage of selection and no-grouping (SNG) were reported instead of GA and SGA, respectively. In terms of prediction and selection, the QR-PACS method does not perform well, while the QR-elastic-net, QR-adaptive Lasso, and QR-SCAD perform the best, respectively. All the methods under consideration perform well in terms of not identifying the group. Thus, the QR-PACS is not a recommended method when there is high correlation and the significant variables do not form a group.

Table 3 demonstrates that the QR-elastic-net and QR-adaptive Lasso have the best SA, respectively; however, the QR-PACS performs better in terms of ME. It is obvious that QR-PACS identifies the important group with high GA and SGA.

Table 4 shows that the QR-elastic-net, QR-adaptive Lasso, and QR-SCAD have the best SA. It is clear that the QR-PACS has the best results among the other methods in terms of ME. In terms of GA and SGA, it can be observed that the QR-PACS performs well.

From Table 5, it can be noticed that the QR-elastic-net has the best SA. The QR-PACS has excellent GA. Also, it is clear that QR-PACS successfully identifies the groups of predictors as seen in the GA and SGA.

5. NCAA Sports Data

In this section, the behavior of the QR-PACS with ridge QR, QR-Lasso, QR-SCAD, QR-adaptive Lasso, and QR-elastic-net was illustrated in the analysis of NCAA sports data [21]. We standardized the predictors and centered the response before the data analysis.

In each repetition, the authors randomly split the data into a training and a testing dataset, the percentage of the testing data was 20%, and models were fit onto the training set. The NCAA sports data were randomly split 100 times each to allow for more stable comparisons. We reported the average and SE of the ratio of test error (RTE) over QR of all methods and the effective model size (MZ) after accounting for equality of absolute coefficient estimates.

The NCAA data was taken from a study of the effects of sociodemographic indicators and the sports programs on graduation rates. The data size is n=94 and p=19 predictors. The dataset and its description are available from the website (http://www4.stat.ncsu.edu/~boos/var.select/ncaa.html). The predictors are students in top 10% HS (x₁), ACT COMPOSITE 25TH (x₂), On living campus (x₃), first-time undergraduates (x₄), Total Enrolment/1000 (x₅), courses taught by TAs (x₆), composite of basketball ranking (x₇), in-state tuition/1000 (x₈), room and board/1000 (x₉), avg BB home attendance (x₁₀), Professor Salary (x₁₁), Ratio of Student/faculty (x₁₂), white (x₁₃), Assistant professor salary (x₁₄), population of city (x₁₅), faculty with PHD (x₁₆), Acceptance rate (x₁₇), receiving loans (x₁₈), and Out of state (x₁₉).
### Table 3: ME, SA, GA, and SGA results of Example 3 for n=100.

| \( \tau \) | Criterion | Ridge | QR-LASSO | QR-SCAD | QR-adaptive LASSO | QR-elastic-net | QR-PACS |
|---|---|---|---|---|---|---|---|
| 0.10 | ME(S.E) | 0.1383(0.0195) | 0.1176(0.0196) | 0.0933(0.0164) | 0.0908(0.0151) | 0.0902(0.0119) | 0.0558(0.0098) |
| | SA | 0.0 | 58 | 79 | 79 | 82 | 78 |
| | GA | 0.0 | 0 | 0 | 0 | 0 | 85 |
| | SGA | 0.0 | 0 | 0 | 0 | 0 | 70 |
| 0.50 | ME(S.E) | 0.1139(0.0071) | 0.0843(0.0062) | 0.0662(0.0059) | 0.0643(0.0054) | 0.0638(0.0051) | 0.0367(0.0029) |
| | SA | 0.0 | 62 | 80 | 81 | 84 | 79 |
| | GA | 0.0 | 0 | 0 | 0 | 0 | 86 |
| | SGA | 0.0 | 0 | 0 | 0 | 0 | 71 |
| 0.75 | ME(S.E) | 0.1289(0.0180) | 0.1083(0.0171) | 0.0861(0.0161) | 0.0843(0.0147) | 0.0834(0.0108) | 0.0479(0.0081) |
| | SA | 0.0 | 61 | 80 | 80 | 84 | 79 |
| | GA | 0.0 | 0 | 0 | 0 | 0 | 86 |
| | SGA | 0.0 | 0 | 0 | 0 | 0 | 71 |

### Table 4: ME, SA, GA, and SGA results of Example 4 for n=100.

| \( \sigma \) | Criterion | Ridge | QR-LASSO | QR-SCAD | QR-adaptive LASSO | QR-elastic-net | QR-PACS |
|---|---|---|---|---|---|---|---|
| 0.10 | ME(S.E) | 0.1490(0.0199) | 0.1279(0.0198) | 0.1030(0.0169) | 0.1016(0.0155) | 0.0998(0.0124) | 0.0651(0.0104) |
| | SA | 0.0 | 57 | 78 | 78 | 80 | 77 |
| | GA | 0.0 | 0 | 0 | 0 | 0 | 83 |
| | SGA | 0.0 | 0 | 0 | 0 | 0 | 69 |
| 0.50 | ME(S.E) | 0.1244(0.0076) | 0.0945(0.0065) | 0.0867(0.0062) | 0.0742(0.0056) | 0.0735(0.0054) | 0.0464(0.0032) |
| | SA | 0.0 | 61 | 79 | 80 | 83 | 78 |
| | GA | 0.0 | 0 | 0 | 0 | 0 | 85 |
| | SGA | 0.0 | 0 | 0 | 0 | 0 | 70 |
| 0.75 | ME(S.E) | 0.1399(0.0185) | 0.1189(0.0171) | 0.0958(0.0166) | 0.0946(0.0160) | 0.0930(0.0111) | 0.0567(0.0086) |
| | SA | 0.0 | 60 | 79 | 80 | 83 | 78 |
| | GA | 0.0 | 0 | 0 | 0 | 0 | 85 |
| | SGA | 0.0 | 0 | 0 | 0 | 0 | 70 |
Table 5: ME, SA, GA, and SGA results of Example 5 for n=100.

| \( \tau \) | Criterion | Ridge QR | QR-LASSO | QR-SCAD | QR-adaptive LASSO | QR-elastic-net | QR-PACS |
|---|---|---|---|---|---|---|---|
| 0.10 | ME(S.E) | 0.1330(0.0161) | 0.1106(0.0168) | 0.0914(0.0154) | 0.0910(0.0142) | 0.0818(0.0122) | 0.0548(0.0981) |
| | SA | 0 | 58 | 79 | 79 | 84 | 79 |
| | GA | 0 | 0 | 0 | 0 | 0 | 0 |
| | SGA | 0 | 0 | 0 | 0 | 0 | 0 |
| \( \sigma = 1 \) | ME(S.E) | 0.1119(0.0061) | 0.0807(0.0053) | 0.0666(0.0054) | 0.0663(0.0051) | 0.0561(0.0054) | 0.0358(0.0031) |
| 0.50 | SA | 0 | 63 | 82 | 82 | 86 | 81 |
| | GA | 0 | 0 | 0 | 0 | 0 | 0 |
| | SGA | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.75 | SA | 0 | 62 | 81 | 81 | 84 | 73 |
| | GA | 0 | 0 | 0 | 0 | 0 | 0 |
| | SGA | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.10 | ME(S.E) | 0.1486(0.0188) | 0.1260(0.0189) | 0.1001(0.0156) | 0.1000(0.0148) | 0.0960(0.0110) | 0.0625(0.0088) |
| 0.50 | SA | 0 | 57 | 77 | 77 | 81 | 78 |
| | GA | 0 | 0 | 0 | 0 | 0 | 0 |
| | SGA | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.75 | SA | 0 | 63 | 80 | 80 | 83 | 80 |
| | GA | 0 | 0 | 0 | 0 | 0 | 0 |
| | SGA | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.10 | ME(S.E) | 0.1274(0.0169) | 0.1048(0.0162) | 0.0867(0.0157) | 0.0865(0.0143) | 0.0756(0.0109) | 0.0476(0.0079) |
| 0.50 | SA | 0 | 62 | 81 | 81 | 84 | 79 |
| | GA | 0 | 0 | 0 | 0 | 0 | 0 |
| | SGA | 0 | 0 | 0 | 0 | 0 | 0 |
| 0.75 | SA | 0 | 60 | 80 | 80 | 83 | 79 |
| | GA | 0 | 0 | 0 | 0 | 0 | 0 |
| | SGA | 0 | 0 | 0 | 0 | 0 | 0 |

QR-adaptive Lasso, and QR-elastic-net in test error. In fact, ridge QR, QR-Lasso, QR-SCAD, QR-adaptive Lasso, and QR-elastic-net perform worse than the QR in test error. The effective model size is 5 for QR-PACS, although it includes all variables in the models.

6. Conclusions

In this paper, QR-PACS for group identification and VS under QR settings is developed, which combines the strength of QR and the ability of PACS for consistent group identification and VS. QR-PACS can achieve the two goals simultaneously. QR-PACS extends PACS from mean regression settings to QR settings. It is proved computationally that it can be simply carried out with an effective computational algorithm. The paper shows that QR-PACS can yield promising predictive precision as well as identifying related groups in both simulation and the real data. Future direction or extension of the current paper is QR-PACS under Bayesian framework. Also, robust QR-PACS is another possible extension of the current paper.

Data Availability

The data which is studied in our paper is the NCAA sports data from Mangold et al. [21]. It is public and available from the website (http://www4.stat.ncsu.edu/~boos/var.select/ncaa.html), [21].
Conflicts of Interest
The authors declare that there are no conflicts of interest regarding the publication of this paper.

References
[1] R. Koenker and G. Bassett Jr., “Regression quantiles,” *Econometrica*, vol. 46, no. 1, pp. 33–50, 1978.
[2] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 58, no. 1, pp. 267–288, 1996.
[3] J. Fan and R. Li, “Variable selection via nonconcave penalized likelihood and its oracle properties,” *Journal of the American Statistical Association*, vol. 96, no. 456, pp. 1348–1360, 2001.
[4] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, and K. Knight, “Sparsity and smoothness via the fused lasso,” *Journal of the Royal Statistical Society B: Statistical Methodology*, vol. 67, no. 1, pp. 91–108, 2005.
[5] H. Zou and T. Hastie, “Regularization and variable selection via the elastic net,” *Journal of the Royal Statistical Society B: Statistical Methodology*, vol. 67, no. 2, pp. 301–320, 2005.
[6] M. Yuan and Y. Lin, “Model selection and estimation in regression with grouped variables,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 1, pp. 49–67, 2006.
[7] H. Zou, “The adaptive lasso and its oracle properties,” *Journal of the American Statistical Association*, vol. 101, no. 476, pp. 1418–1429, 2006.
[8] H. Zou and H. H. Zhang, “On the adaptive elastic-net with a diverging number of parameters,” *The Annals of Statistics*, vol. 37, no. 4, pp. 1733–1751, 2009.
[9] C. H. Zhang, “Nearly unbiased variable selection under minimax concave penalty,” *The Annals of Statistics*, vol. 38, no. 2, pp. 894–942, 2010.
[10] R. Koenker, “Quantile regression for longitudinal data,” *Journal of Multivariate Analysis*, vol. 91, no. 1, pp. 74–89, 2004.
[11] H. Wang, G. Li, and G. Jiang, “Robust regression shrinkage and consistent variable selection through the LAD-Lasso,” *Journal of Business & Economic Statistics*, vol. 25, no. 3, pp. 347–355, 2007.
[12] Y. Li and J. Zhu, “$L_1$-norm quantile regression,” *Journal of Computational and Graphical Statistics*, vol. 17, no. 1, pp. 163–185, 2008.
[13] Y. Wu and Y. Liu, “Variable selection in quantile regression,” *Statistica Sinica*, vol. 19, no. 2, pp. 801–817, 2009.
[14] M. Slawski, “The structured elastic net for quantile regression and support vector classification,” *Statistics and Computing*, vol. 22, no. 1, pp. 153–168, 2012.
[15] A. Belloni and V. Chernozhukov, $L_1$–Penalized Quantile Regression in High-Dimensional Sparse Models, 2011.
[16] L. Wang, Y. Wu, and R. Li, “Quantile regression for analyzing heterogeneity in ultra-high dimension,” *Journal of the American Statistical Association*, vol. 107, no. 497, pp. 214–222, 2012.
[17] B. Peng and L. Wang, “An iterative coordinate descent algorithm for high-dimensional nonconvex penalized quantile regression,” *Journal of Computational and Graphical Statistics*, vol. 24, no. 3, pp. 676–694, 2015.
[18] D. B. Sharma, H. D. Bondell, and H. H. Zhang, “Consistent group identification and variable selection in regression with correlated predictors,” *Journal of Computational and Graphical Statistics*, vol. 22, no. 2, pp. 319–340, 2013.
Submit your manuscripts at
www.hindawi.com