Journal of Mathematics and Statistics

Original Research Paper

Sparse Sliced Inverse Quantile Regression

Ali Alkenani and Tahir R. Dikheel

Department of Statistics, College of Administration and Economics, University of Al-Qadisiyah, Iraq

Article history
Received: 05-02-2016
Revised: 08-09-2016
Accepted: 10-09-2016

Corresponding Author:
Ali Alkenani
Department of Statistics,
College of Administration and Economics, University of Al-Qadisiyah, Iraq
Email: ali.alkenani@qu.edu.iq

Abstract: The current paper proposes the sliced inverse quantile regression method (SIQR). In addition to the latter this study proposes both the sparse sliced inverse quantile regression method with Lasso (LSIQR) and Adaptive Lasso (ALSIQR) penalties. This article introduces a comprehensive study of SIQR and sparse SIQR. The simulation and real data analysis have been employed to check the performance of the SIQR, LSIQR and ALSIQR. According to the results of median of mean squared error and the absolute correlation criteria, we can conclude that the SIQR, LSIQR and ALSIQR are the more advantageous approaches in practice.

Keywords: Dimension Reduction, Variable Selection, Sliced Inverse Quantile Regression, Lasso, Adaptive Lasso

Introduction

In many statistical applications, the number of variables becomes huge. Consequently, many of statistical data analyses become hard. A familiar way to cope with this issue is to shrinkage the dimension of the regression model, without much loss of information on regression. This has been obtained via the Sufficient Dimension Reduction (SDR) theory.

The SDR theory (Cook, 1998) aims to reduce the p-dimensional predictor X with a d-dimensional vector $\beta^T X$, where β is a $p\times d$ matrix with $d\leq p$, without much loss of information on regression and putting only a few assumptions. The subspace spanned by the columns of β, Span (β), is called a Dimension Reduction Subspace (D.R.S). The minimal subspace is usually uniquely defined in practice and coincides with the intersection of all of the subspaces (Cook, 1996). Such an intersection is a parsimonious population parameter that contains all regression information of Y given X and thus it is the central matter of concern in Dimension Reduction (D.R). This intersection is called the Central Dimension Reduction (CDR) space and is written as $S_{Y|X}$ and its dimension, $d = dim\ (S_{Y|X})$, is called the structural dimension of regression (Cook, 1998).

There have been a number of methods suggested to find the SDR in regression through estimating the Central Subspace (C.S) and one of the well-known methods for estimating C.S is Sliced Inverse Regression (SIR) (Li, 1991). The SIR is especially practical in dealing with high-dimensional covariates, and it has been shown to be an powerful D.R tool in high-dimensional regression problems (Zhu *et al*., 2006). Li (1991) suggested that an estimate of the $S_{Y|X}$ can be achieved by the first d eigenvectors $\upsilon_1,\ldots,\upsilon_d$ for the eigenvalue problem of the form:

$$M\upsilon_i = \rho_i \Sigma_x \upsilon_i, \tag{1}$$

where, $\rho_1\geq\cdots\geq \rho_d>0$ are the eigenvalues, $\Sigma_x = Cov(X)$ and $M = Cov\{E(X|Y)\}$.

Aragon and Saracco (1997) studied the finite sample properties of SIR. The Lasso has been combined with SIR in (Ni *et al*., 2005; Li and Nachtsheim, 2006) to produce sparse estimates. Li (2007) proposed to combine a regression-type formulation of SDR methods with shrinkage estimation to produce sparse and accurate solutions. This strategy can be applied to SIR and many of SDR methods.

Li and Yin (2008) suggested a penalized SIR based on the Least-Squares (L.S) formulation of SIR. Cook (2004) rewrote SIR in (1) as a L.S minimization problem and SIR estimates can be obtained by minimizing:

$$L(B,C) = \sum_{y=1}^{h} \hat{f}_y \left\| \bar{Z}_y - BC_y \right\|^2 \tag{2}$$

where, $\hat{Z} = \hat{\Sigma}^{-\frac{1}{2}}(X - E(X))$ with $\bar{Z}_y$ denoting the mean of $\hat{Z}$ in the y_{th} slice, n_y is the number of observations within each slice and $\hat{f}_y = n_y / n$ is the observed fraction of observations in slice y and h is the number of non-overlapping slices. Over $B\in R^{p\times d}$ and $C = (C_1,\ldots,C_h)\in R^{d\times h}$ the values of B which minimize $L(B,C)$ form an estimation

Science Publications

© 2016 Ali Alkenani and Tahir R. Dikheel. This open access article is distributed under a Creative Commons Attribution (CC-BY) 3.0 license.

of the central space $S_{Y|X}$. Equation 2 requires $\hat{\Sigma}^{-\frac{1}{2}}$. The inversion is not possible and a penalization approach has to be used in case of high correlations between the predictors or small sample sizes compared to the dimension. The L.S formulation of SIR in the original predictor X scale has been derived in (Li and Yin, 2008) in order to avoid the singularity of $\hat{\Sigma}_x$ as follows:

$$\tilde{L}(B,C)=\sum_{y=1}^{h}\hat{f}_y\left\|\left(\bar{X}_y-\bar{X}\right)-\hat{\Sigma}_x BC_y\right\|^2 \quad (3)$$

The mechanics of the alternating L.S algorithm which is suggested by Li and Yin (2008) to minimize (3) can be described as follow.

Given B the solution of C can be obtained by:

$$\hat{C}=\left(\hat{C}_1,\ldots,\hat{C}_h\right) \text{ where } \hat{C}_y=\left(B^T\hat{\Sigma}_x^2B\right)^{-1}B^T\hat{\Sigma}_x\left(\bar{X}_y-\bar{X}\right) \quad (4)$$

Thereafter, rewrite (3) in the following form:

$$\tilde{L}(B,C)=\left\|\tilde{A}^{1/2}\tilde{Y}-\tilde{A}^{1/2}\left(C^T\otimes\hat{\Sigma}_x\right)vec(B)\right\|^2 \quad (5)$$

where, $vec(.)$ is a matrix operator.

$\otimes$ is the Kronecker product, $\tilde{Y}=vec\left(\bar{X}_1-\bar{X},\ldots,\bar{X}_h-\bar{X}\right)$, $\tilde{A}^{1/2}=D_f^{1/2}\otimes I_P$ and $D_f=diag\left(\hat{f}_1,\ldots,\hat{f}_h\right)$. Given C, the solution of B in (5) is:

$$vec\left(\hat{B}\right)=(CD_fC^T\otimes\hat{\Sigma}_x^2)^{-1}\left(CD_f\otimes\hat{\Sigma}_x\right)\tilde{Y} \quad (6)$$

and this procedure will continue between minimizing B and C until convergence:

$$\tilde{L}\left(\hat{B},\hat{C}\right)=\arg\min_{B,C}\left\|\tilde{A}^{1/2}\tilde{Y}-\tilde{A}^{1/2}\left(C^T\otimes\Sigma_x\right)vec(B)\right\|^2 \quad (7)$$

where, $\left(\hat{B},\hat{C}\right)$ denote the SIR estimator that minimizes (7).

Also, Li and Yin (2008) proposed shrinkage SIR estimator of $S_{Y|X}$ as $Span\left(diag(\hat{\alpha})\hat{B}\right)$, where $\hat{\alpha}=\left(\hat{\alpha}_1,\ldots,\hat{\alpha}_P\right)^T\in\mathbb{R}^p$ is obtained by solving:

$$\arg\min_{\alpha}\sum_{y=1}^{h}\hat{f}_y\left\|\left(\bar{X}_y-\bar{X}\right)-\hat{\Sigma}_x diag\left(\hat{B}\hat{C}_y\right)\alpha\right\|^2+\lambda\sum_{j=1}^{P}\left|\alpha_j\right| \quad (8)$$

where, $\lambda > 0$ is the penalty tuning parameter.

The authors wrote:

$$\begin{aligned}&\tilde{Y}=vec\left(\bar{X}_1-\bar{X},\ldots,\bar{X}_h-\bar{X}\right)\\&\in\mathbb{R}^{ph},\tilde{X}=(diag\left(\hat{B}\hat{C}_1\right)\hat{\Sigma}_x,\ldots,diag\left(\hat{B}\hat{C}_y\right)\hat{\Sigma}_x)^T\in\mathbb{R}^{ph\times p}\end{aligned} \quad (9)$$

Then α is the Lasso estimator for the regression of $\tilde{Y}$ on the p-dimensional "data matrix" $\tilde{X}$.

Quantile Regression (QR) has become well-known approach to describe the distribution of a response variable given a set of predictors. QR gives a complete analysis of the stochastic relationships among random variables. The QR has been used in different areas such as finance, microarrays and many other fields (Yu *et al*., 2003). While QR has become very attractive as a complete extension of the mean regression; however, it suffers from the 'curse of dimensionality' (C.D).

There are a number of approaches tried to shorten the dimension and then estimate the Conditional Quantile (C.Q). For example, Chaudhuri (1991; Horowitz and Lee, 2005; Dette and Scheder, 2011; Yebin *et al*., 2011). Wu *et al*. (2010) suggested modelling the conditional quantile by a single-index function to tackle the dimensionality problem. Alkenani and Yu (2013) proposed penalized single index QR to reduce the dimensionality.

Gannoun *et al*. (2004) used SIR to tackle the dimensionality problem of the predictors in order to obtain a more efficient estimator of C.Q. Specifically, the authors employ SIR method as a pre-step to avoid C.R. and then conditional quantile estimators are obtained by inverting the conditional distribution.

In this study, one step sliced Inverse Quantile Regression (SIQR) has been proposed, which will inherit the same advantages as in the SIR. In addition, sparse sliced inverse quantile with Lasso and Adaptive Lasso penalties have been suggested. This paper is arranged as follows. The SIQR is proposed in section 2. The LSIQR and ALSIQR are suggested in sections 3 and 4, respectively. Simulation examples and real data are presented in section 5 and 6, respectively. The conclusions are summarized in section 7.

SIQR

We can write the equation (7) as follows:

$$\tilde{L}\left(\hat{B},\hat{C}\right)=\arg\min_{B,C}\left\|Y^*-X^*\beta^*\right\|^2 \quad (10)$$

Where:

$Y^* = \tilde{A}^{1/2}\tilde{Y}$, $X^*=\tilde{A}^{1/2}\left(C^T\otimes\hat{\Sigma}_x\right)$, $\beta^* = vec(B)$

$Y^* = n^*\times 1$ response vector, $n^* = (ph)$

$X^* = n^*\times pd$ predictors matrix

$\beta^* = pd\times 1$ coefficients vector

Then, we can propose SIQR as follows:

$$\widehat{\beta^*}=\arg\min_{B,C}\sum_{i=1}^{n^*}\rho_\tau\left(Y_i^*-X_i^*\beta^*\right) \quad (11)$$

where, ρ_τ (.) is the check function defined by $\rho_\tau(u) = \tau u I_{[0,\infty)}(u) - (1-\tau) u I_{(-\infty,0)}(u)$.

Then we will replace B by $\widehat{\beta^*}$ in Equation 4 to compute $\hat{C}$. The new values of $\hat{C}$ will be put in Equation 11 to compute a new $\widehat{\beta^*}$. This procedure will continue between minimizing B and C until convergence.

The algorithm has been summarized as follows:

- Initialization step) Obtain the initial β_0 from SIR methods where β_0 is $p\times1$ estimated vector coefficients. Set $B = \beta_0$
- Given $B = \beta_0$ find $\hat{C}$ from equation (4)
- Given $\hat{C}$ find $X^* = \tilde{A}^{1/2}\left(C^T \otimes \hat{\Sigma}_x\right)$ where $\tilde{A}^{1/2}$ is defined in Equation 5
- Now, we have X^* and Y^*. $\widehat{\beta^*}$ can be obtained by solving (11) as standard linear QR. We can use *rq*(Y^*~X^*,*tau*, method = "fn") function in quantreg package to find $\widehat{\beta^*}$
- Set $B = \widehat{\beta^*}$ and repeat steps 2,3 and 4 until convergence is attained

LSIQR

Tibshirani (1996), Lasso has been proposed for simultaneous variable selection and parameter estimation. It minimizes the residual sum of squares with a constraint on the l_1 norm. Li and Zhu (2008) extended Lasso Tibshirani (1996) to work with QR models.

From Equation 8, we can propose LSIQR as follow:

$$\arg\min_\alpha \sum_{j=1}^{p}\sum_{y=1}^{h} \hat{f}_y \rho_\tau\left(\tilde{Y}_{yj} - \tilde{X}_{yj}\alpha\right) + \lambda \sum_{j=1}^{p}\left|\alpha_j\right| \quad (12)$$

where, $\tilde{Y}$ and $\tilde{X}$ were defined in equation (9).

The optimization problem in (12) has been solved by employing a standard Lasso QR.

The algorithm has been summarized as follows:

- Let $\hat{B}$ and $\hat{C}$ represent the convergent values for $\hat{B}$ and $\hat{C}$ which we obtained from the previous algorithm
- Given $\hat{B}$ and $\hat{C}$ find $\tilde{X} = (diag\left(\hat{B}\hat{C}_1\right)\hat{\Sigma}_x,\ldots,diag\left(\hat{B}\hat{C}_y\right)\hat{\Sigma}_x)^T \in \mathbb{R}^{ph\times p}$
- Now, we have $\tilde{X}$ and $\tilde{Y}$. Where $\tilde{Y} = vec\left(\bar{X}_1 - \bar{X},\ldots,\bar{X}_h - \bar{X}\right) \in \mathbb{R}^{ph}$. $\hat{\alpha}$ can be obtained by solving (12) as Lasso linear QR. We can use $rq(\tilde{Y} \sim \tilde{X}, tau, \text{method} = \text{"lasso"})$ function in quantreg package to find $\hat{\alpha}$

ALSIQR

Fan and Li (2001) proved that Lasso produces biased estimates and the oracle properties do not hold for the lasso. Adaptive Lasso method, in which adaptive weights are used for penalizing different coefficients in the l_1 penalty, has been suggested in (Zou, 2006). The Adaptive Lasso benefits from the oracle properties that the Lasso does not have (Zou, 2006). The Adaptive Lasso QR has been suggested in (Wu and Liu, 2009).

ALSIQR has been proposed as follows:

$$\arg\min_\alpha \sum_{j=1}^{p}\sum_{y=1}^{h} \hat{f}_y \rho_\tau\left(\tilde{Y}_{yj} - \tilde{X}_{yj}\alpha\right) + \lambda_n \sum_{j=1}^{p} \tilde{w}_j\left|\alpha_j\right| \quad (13)$$

where, the weights are set to be $\tilde{w}_j = \frac{1}{\left|\tilde{\alpha}_j\right|^\gamma}$, $j = 1,\ldots,p$; for some appropriately chosen $\gamma>0$, $\tilde{\alpha}$ is the quantile sliced inverse regression estimates.

We can summarize the algorithm as follows:

- Let $\hat{B}$ and $\hat{C}$ represent the convergent values for $\hat{B}$ and $\hat{C}$ which we obtained from the previous algorithm
- Given $\hat{B}$ and $\hat{C}$ find $\tilde{X} = (diag\left(\tilde{B}\hat{C}_1\right)\hat{\Sigma}_x,\ldots,diag\left(\tilde{B}\hat{C}_y\right)\hat{\Sigma}_x)^T \in \mathbb{R}^{ph\times p}$
- Now, we have $\tilde{X}$ and $\tilde{Y}$. Where $\tilde{Y} = vec\left(\bar{X}_1 - \bar{X},\ldots,\bar{X}_h - \bar{X}\right) \in \mathbb{R}^{ph}$. $\tilde{\alpha}$ can be obtained by solving (13) as Adaptive Lasso linear QR

The LARS algorithm (Efron *et al*., 2004; Zou, 2006) has been applied to get the Adaptive Lasso estimate of $\tilde{\alpha}$ in (13), which is described as follows:

Step 1. $\tilde{X}_{yj}$ is the jth coordinate of $\tilde{X}_y$. For any given λ, define $X^{**}_{yj} = \tilde{X}_{yj} / \tilde{w}_j$, $y = 1,\ldots,h$ and $j = 1,\ldots, p$; where, X^{**} is the re-scaled predictor matrix.

Step 2. Obtain $\widehat{\alpha^*}$ by solving the standard lasso QR problem for all λ_n by using LARS as follows:

$$\widehat{\alpha^*} = \arg\min_\alpha \sum_{j=1}^{p}\sum_{y=1}^{h} \hat{f}_y \rho_\tau\left(\tilde{Y}_{yj} - X^{**}_{yj}\alpha\right) + \lambda \sum_{j=1}^{p}\left|\alpha_j\right| \quad (14)$$

The minimization problem in (14) can be solve by using $rq(\tilde{Y} \sim X^{**}, tau, \text{method} = \text{"lasso"})$ function in quantreg package to find $\widehat{\alpha^*}$.

Step 3. Output $\hat{\alpha}_j = \widehat{\alpha^*_j} / \tilde{w}_j$.

Simulation Study

A-The performance of SIQR has been checked via a numerical study. SIQR estimators have compared with linear QR (LQR) and nonlinear QR (NQR) estimators for τ = (0.10, 0.25, 0.50). To make comparisons, the mean and Standard Deviation (SD) of the absolute correlation (|r|) between the estimated predictor $\hat{\beta}^T X$ and the true predictor $\beta^T X$ and the Median of Mean Squared Errors (MMSE) for $\hat{\beta}^T X$ have been reported.

Example 1: R = 500 samples with size n = 400 observations from $y = \beta^T X + \sigma\varepsilon$ have been generated, where $\beta = (1,1,1,1,1,1,1,1,1,0)^T$, $X \in R^{10}$ and x_i (i = 1,…,10) and ε are independently and identically distributed (i.i.d) standard normal. We take σ= 1 and σ= 3.

Example 2: R = 200 data-sets with size n = 400 observations have been generated from the following model:

$$y = \sin\left\{\frac{\pi(u-A)}{C-A}\right\} + \varepsilon$$

where, $u = \beta^T X$, $X = (x_1,\ldots,x_8)$, $\beta = (1,1,1,1,1,1,1,1)^T / \sqrt{8}$, $A = \frac{\sqrt{3}}{2} - \frac{1.645}{\sqrt{12}}$, $C = \frac{\sqrt{3}}{2} + \frac{1.645}{\sqrt{12}}$, x_i i.i.d.~ Unif(0,1), i = 1,2,…,8; $\varepsilon \sim N(0,1)$; x_i' s and ε are i.i.d. β is estimated with τ = (0.10, 0.25, 0.50).

B- In term of variable selection, the performance of LSIQR and ALSIQR has been tested through a numerical study. The LSIQR and ALSIQR have compared with Lasso linear QR (LLQR) and Adaptive Lasso linear QR (ALLQR) for τ = (0.10, 0.25, 0.50). The average number of zero coefficients (Ave 0's), the mean and SD of |r| between $\hat{\beta}^T X$ and $\beta^T X$ and MMSE for $\hat{\beta}^T X$ have been reported. $\hat{\beta}$ is assumed zero if $|\hat{\beta}|$ is smaller than 10^{-6}.

Example 3: R = 500 samples have been generated with n = 400 from = $\beta^T X + \sigma\varepsilon$, where β takes the following different forms:

Model 1: $\beta = (1,1,0.1,0.1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)^T$
Model 2: $\beta = (1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)^T$
Model 3: $\beta = (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0.1,0.1,)^T$
Model 4: $\beta = (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1)^{\mathrm{T}}$
Model 5: $\beta = (1,1,0.1,0.1,0,0,0,0,0,0,0,0,0,0,0,0,1,1,0.1,0.1)^T$
Model 6: $\beta = (1,1,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,1,1)^T$

$X \in R^{20}$ and X = $(x_1,\ldots,x_{20})$ are generated from $N(0,\Sigma)$ and the (i,j) element of Σ is $0.5^{|i-j|}$. The error term ε is generated from standard normal distribution. We take σ = 1 and σ= 3.

Example 4: R = 500 samples has been generated with n = 400 from $y = exp(-0.5\ \beta^T X) + 0.2\varepsilon$, where X = $(x_1,\ldots,x_6)$, and X and the error ε are independent standard normal variables. The C.S is spanned by $\beta = (1,-1,0,0,0,0)^T / \sqrt{2}$. β is estimated with τ= (0.10, 0.25, 0.50).

According to the mean and SD of |r_i| between $\hat{\beta}_j^T X$ and $\beta_j^T X$ and MMSE for $\hat{\beta}^T X$ (From Table 1-9), it can be seen that ALSIQR, LSIQR and SIQR have a better performance than the LLQR, ALLQR, LQR and NQR for all studied cases. It is obvious from the Table 1 and 2 the preference of SIQR, when it compares with LQR and NQR, depending on MMSE for both σ values and all values of τ. Furthermore, when the values of τ go up the values of MMSE go down.

From Table 3-8 we find that ALSIQR and LSIQR give MMSE and SD values less than the other methods. Also, the results show that the MMSE values for the all considered methods increase when σ = 1 move to σ = 3 for all τ values. Moreover, the values of MMSE for all methods increase when the values of τ decrease.

Table 1. Simulation results for the SIQR and LQR based on the linear model in example 1

		τ = 0.10		τ = 0.25		τ = 0.50	
		LQR	SIQR	LQR	SIQR	LQR	SIQR
σ= 1	Mean |r|	0.9915	0.9966	0.9920	0.9977	0.9924	0.9981
	SD |r|	0.0002	0.0001	0.0002	0.0001	0.0002	0.0001
	MMSE	0.0007	0.0005	0.0005	0.0002	0.0006	0.0001
σ= 3	Mean |r|	0.9700	0.9753	0.9768	0.9785	0.9764	0.9823
	SD |r|	0.0003	0.0001	0.0002	0.0001	0.0002	0.0001
	MMSE	0.0044	0.0014	0.0014	0.0010	0.0020	0.0002

Table 2. Simulation results for the SIQR and NQR based on the nonlinear model in example 2

	τ = 0.10		τ = 0.25		τ = 0.50	
	NQR	SIQR	NQR	SIQR	NQR	SIQR
Mean |r|	0.8822	0.8911	0.8740	0.8966	0.9037	0.9201
SD |r|	0.1048	0.0889	0.0940	0.0863	0.0627	0.0450
MMSE	0.0067	0.0060	0.0061	0.0059	0.0063	0.0055

Table 3. Simulation results for the ALSIQR, LSIQR, SIQR, ALLQR, LLQR and LQR based on the linear model in example 3 model 1, $\sigma = 1$ and $\sigma = 3$

		LQR	LLQR	ALLQR	SIQR	LSIQR	ALSIQR
$\tau = 0.10$							
	Ave 0's	2.4300	8.7700	13.1300	3.6300	10.2700	13.1300
$\sigma = 1$	Mean $\|r\|$	0.9829	0.9839	0.9892	0.9693	0.9925	0.9952
	SD $\|r\|$	0.0071	0.0073	0.0055	0.0101	0.0022	0.0021
	MMSE	0.0027	0.0025	0.0020	0.0018	0.0017	0.0010
$\sigma = 3$	Ave 0's	9.3000	9.8500	11.8000	9.4500	9.6000	11.4000
	Mean $\|r\|$	0.8518	0.8600	0.8690	0.8855	0.9401	0.9552
	SD $\|r\|$	0.0686	0.0675	0.0673	0.0301	0.0209	0.0171
	MMSE	0.0237	0.0205	0.0206	0.0028	0.0020	0.0024
$\tau = 0.25$							
$\sigma = 1$	Ave 0's	2.4500	8.7500	13.2000	2.9500	9.9000	12.6500
	Mean $\|r\|$	0.9860	0.9871	0.9921	0.9683	0.9923	0.9947
	SD $\|r\|$	0.0037	0.0037	0.0035	0.0111	0.0022	0.0017
	MMSE	0.0030	0.0029	0.0027	0.0029	0.0026	0.0023
$\sigma = 3$	Ave 0's	3.4000	8.8000	12.0500	3.8500	10.2000	11.5500
	Mean $\|r\|$	0.9118	0.9178	0.9295	0.8900	0.9445	0.9584
	SD $\|r\|$	0.0343	0.0313	0.0295	0.0312	0.0158	0.0152
	MMSE	0.0053	0.0051	0.0039	0.0037	0.0027	0.0025
$\tau = 0.50$							
	Ave 0's	3.1000	8.4000	13.3000	3.3000	9.9000	12.8000
$\sigma = 1$	Mean $\|r\|$	0.9890	0.9900	0.9942	0.9691	0.9944	0.9946
	SD $\|r\|$	0.0033	0.0031	0.0030	0.0118	0.0025	0.0022
	MMSE	0.0007	0.0005	0.0002	0.0010	0.0002	0.0001
$\sigma = 3$	Ave 0's	8.5500	8.9500	12.3000	10.1000	9.5500	11.7500
	Mean $\|r\|$	0.9119	0.9194	0.9351	0.8989	0.9486	0.9632
	SD $\|r\|$	0.0266	0.0264	0.0259	0.0389	0.0193	0.0180
	MMSE	0.0009	0.0007	0.0007	0.0014	0.0005	0.0005

Table 4. Simulation results for the ALSIQR, LSIQR, SIQR, ALLQR, LLQR and LQR based on the linear model in example 3 model 2, $\sigma = 1$ and $\sigma = 3$

		LQR	LLQR	ALLQR	SIQR	LSIQR	ALSIQR
$\tau = 0.10$							
$\sigma = 1$	Ave 0's	3.00000	8.15000	12.3500	3.5500	9.6000	12.0500
	Mean $\|r\|$	0.98870	0.98950	0.9930	0.9737	0.9947	0.9967
	SD $\|r\|$	0.00420	0.00420	0.0039	0.0103	0.0019	0.0017
	MMSE	0.00330	0.00280	0.0028	0.0021	0.0018	0.0014
$\sigma = 3$	Ave 0's	8.75000	8.95000	11.1000	11.7500	11.2000	13.1500
	Mean $\|r\|$	0.90860	0.91420	0.9225	0.9276	0.9688	0.9775
	SD $\|r\|$	0.03710	0.03580	0.0373	0.0301	0.0149	0.0118
	MMSE	0.03030	0.02750	0.0226	0.0110	0.0101	0.0103
$\tau = 0.25$							
$\sigma = 1$	Ave 0's	3.45000	8.95000	13.3500	4.5000	9.6000	12.7500
	Mean $\|r\|$	0.99240	0.99290	0.9960	0.9732	0.9962	0.9970
	SD $\|r\|$	0.00330	0.00310	0.0021	0.0138	0.0020	0.0015
	MMSE	0.00070	0.00220	0.0018	0.0031	0.0006	0.0004
$\sigma = 3$	Ave 0's	9.00000	8.90000	11.4500	10.3500	10.3000	11.7500
	Mean $\|r\|$	0.93780	0.94230	0.9531	0.9222	0.9691	0.9786
	SD $\|r\|$	0.01750	0.01670	0.0164	0.0358	0.0116	0.0106
	MMSE	0.00590	0.00530	0.0047	0.0065	0.0045	0.0045
$\tau = 0.50$							
	Ave 0's	4.65000	9.35000	14.4000	4.6000	10.8500	13.0000
$\sigma = 1$	Mean $\|r\|$	0.99440	0.99510	0.9955	0.9781	0.9972	0.9973
	SD $\|r\|$	0.00350	0.00320	0.0030	0.0071	0.0019	0.0016
	MMSE	0.00520	0.00490	0.0004	0.0053	0.0001	0.0001
$\sigma = 3$	Ave 0's	8.40000	8.90000	12.0000	9.7000	9.4500	11.2000
	Mean $\|r\|$	0.94919	0.95312	0.9630	0.9268	0.9683	0.9766
	SD $\|r\|$	0.01321	0.01296	0.0123	0.0332	0.0096	0.0110
	MMSE	0.00260	0.00300	0.0007	0.0025	0.0005	0.0005

Table 5. Simulation results for the ALSIQR, LSIQR, SIQR, ALLQR, LLQR and LQR based on the linear model in example 3 model 3, $\sigma = 1$ and $\sigma = 3$

		LQR	LLQR	ALLQR	SIQR	LSIQR	ALSIQR
$\tau = 0.10$							
$\sigma = 1$	Ave 0's	3.4000	8.5800	12.3000	3.4100	10.3800	13.2800
	Mean $\|r\|$	0.9807	0.9820	0.9878	0.9675	0.9923	0.9946
	SD $\|r\|$	0.0056	0.0059	0.0048	0.0106	0.0033	0.0028
	MMSE	0.0038	0.0033	0.0033	0.0027	0.0011	0.0015
$\sigma = 3$	Ave 0's	8.5800	8.9500	11.3500	9.8300	9.8000	11.6800
	Mean $\|r\|$	0.8662	0.8715	0.8852	0.8835	0.9422	0.9574
	SD $\|r\|$	0.0531	0.0521	0.0500	0.0403	0.0211	0.0204
	MMSE	0.0169	0.0145	0.0128	0.0028	0.0025	0.0015
$\tau = 0.25$							
$\sigma = 1$	Ave 0's	2.4300	8.8300	13.2800	2.8800	9.6800	12.0500
	Mean $\|r\|$	0.9881	0.9888	0.9932	0.9662	0.9934	0.9952
	SD $\|r\|$	0.0036	0.0035	0.0025	0.0136	0.0018	0.0019
	MMSE	0.0010	0.0007	0.0008	0.0006	0.0005	0.0002
$\sigma = 3$	Ave 0's	8.5000	8.8300	11.7000	9.9500	9.6300	11.8800
	Mean $\|r\|$	0.9056	0.9099	0.9217	0.8921	0.9458	0.9585
	SD $\|r\|$	0.0315	0.0306	0.0297	0.0391	0.0195	0.0178
	MMSE	0.0041	0.0040	0.0025	0.0041	0.0021	0.0023
$\tau = 0.50$							
	Ave 0's	2.1300	8.6500	13.3300	2.5300	10.6300	13.0800
$\sigma = 1$	Mean $\|r\|$	0.9887	0.9895	0.9940	0.9692	0.9942	0.9952
	SD $\|r\|$	0.0036	0.0034	0.0025	0.0118	0.0023	0.0019
	MMSE	0.0022	0.0018	0.0018	0.0020	0.0001	0.0001
$\sigma = 3$	Ave 0's	8.6000	9.0800	11.7800	10.2000	9.9500	12.0500
	Mean $\|r\|$	0.9097	0.9160	0.9286	0.8869	0.9435	0.9566
	SD $\|r\|$	0.0280	0.0261	0.0275	0.0356	0.0176	0.0186
	MMSE	0.0024	0.0023	0.0021	0.0023	0.0005	0.0005

Table 6. Simulation results for the ALSIQR, LSIQR, SIQR, ALLQR, LLQR and LQR based on the linear model in example 3 model 4, $\sigma = 1$ and $\sigma = 3$

		LQR	LLQR	ALLQR	SIQR	LSIQR	ALSIQR
$\tau = 0.50$							
$\sigma = 1$	Ave 0's	3.7800	9.2300	12.7800	3.5300	10.2500	12.5800
	Mean $\|r\|$	0.9893	0.9899	0.9932	0.9760	0.9947	0.9967
	SD $\|r\|$	0.0034	0.0033	0.0030	0.0075	0.0015	0.0013
	MMSE	0.0020	0.0019	0.0016	0.0019	0.0015	0.0013
$\sigma = 3$	Ave 0's	8.7300	9.2000	10.9800	10.0000	10.2500	12.1300
	Mean $\|r\|$	0.9070	0.9115	0.9201	0.9113	0.9631	0.9719
	SD $\|r\|$	0.0288	0.0282	0.0270	0.0262	0.0110	0.0100
	MMSE	0.0215	0.0207	0.0184	0.0034	0.0021	0.0024
$\tau = 0.25$							
$\sigma = 1$	Ave 0's	3.5000	8.7300	13.2000	3.8000	9.6800	12.0500
	Mean $\|r\|$	0.9925	0.9931	0.9951	0.9769	0.9961	0.9967
	SD $\|r\|$	0.0023	0.0022	0.0018	0.0073	0.0017	0.0015
	MMSE	0.0004	0.0004	0.0002	0.0004	0.0002	0.0001
$\sigma = 3$	Ave 0's	8.500	8.9500	11.2750	9.6250	9.925	11.5500
	Mean $\|r\|$	0.9367	0.9412	0.9494	0.9204	0.9643	0.9737
	SD $\|r\|$	0.0197	0.0180	0.0180	0.0265	0.0129	0.0107
	MMSE	0.0045	0.0049	0.0036	0.0036	0.0023	0.0023
$\tau = 0.50$							
	Ave 0's	3.3000	8.9500	13.0750	3.3300	10.300	12.4300
$\sigma = 1$	Mean $\|r\|$	0.9933	0.9938	0.9947	0.9737	0.9961	0.9966
	SD $\|r\|$	0.0018	0.0018	0.0017	0.0109	0.0015	0.0013
	MMSE	0.0040	0.0022	0.0002	0.0048	0.0001	0.0001
$\sigma = 3$	Ave 0's	8.8500	9.3500	11.9500	9.7000	10.2300	12.1000
	Mean $\|r\|$	0.9446	0.9485	0.9576	0.9266	0.9651	0.9733
	SD $\|r\|$	0.0173	0.0161	0.0147	0.0298	0.0105	0.0115
	MMSE	0.0055	0.0053	0.0021	0.0065	0.0014	0.0012

Table 7. Simulation results for the ALSIQR, LSIQR, SIQR, ALLQR, LLQR and LQR based on the linear model in example 3 model 5, $\sigma = 1$ and $\sigma = 3$

		LQR	LLQR	ALLQR	SIQR	LSIQR	ALSIQR
$\tau = 0.10$							
$\sigma = 1$	Ave 0's	2.58000	6.73000	10.1300	2.1800	9.4000	11.9000
	Mean $\|r\|$	0.99070	0.99120	0.9934	0.9743	0.9955	0.9958
	SD $\|r\|$	0.00300	0.00300	0.9934	0.0104	0.0018	0.0016
	MMSE	0.00670	0.00670	0.0063	0.0030	0.0029	0.0021
$\sigma = 3$	Ave 0's	7.45000	7.90000	9.7300	10.9300	10.8000	12.9000
	Mean $\|r\|$	0.92380	0.92800	0.9351	0.9318	0.9698	0.9741
	SD $\|r\|$	0.02560	0.02490	0.0252	0.0191	0.0086	0.0131
	MMSE	0.02520	0.02390	0.0192	0.0071	0.0048	0.0059
$\tau = 0.25$							
$\sigma = 1$	Ave 0's	3.23000	6.63000	10.6800	3.3300	10.7300	12.8800
	Mean $\|r\|$	0.99360	0.99400	0.9959	0.9735	0.9960	0.9961
	SD $\|r\|$	0.00210	0.00200	0.0014	0.0101	0.0013	0.0013
	MMSE	0.00060	0.00060	0.0004	0.0005	0.0001	0.0004
$\sigma = 3$	Ave 0's	7.48000	7.95000	10.0800	10.6300	10.6800	12.7300
	Mean $\|r\|$	0.94910	0.95280	0.9576	0.9311	0.9715	0.9767
	SD $\|r\|$	0.01910	0.01850	0.0208	0.0268	0.0120	0.0110
	MMSE	0.00900	0.00780	0.0076	0.0077	0.0075	0.0070
$\tau = 0.50$							
	Ave 0's	2.55000	6.83000	10.7000	2.1800	9.2500	11.7000
$\sigma = 1$	Mean $\|r\|$	0.99390	0.99430	0.9950	0.9755	0.9951	0.9954
	SD $\|r\|$	0.00220	0.00210	0.0021	0.0082	0.0020	0.0020
	MMSE	0.00020	0.00020	0.0002	0.0004	0.0001	0.0001
$\sigma = 3$	Ave 0's	7.37500	7.72500	10.3500	8.8800	8.7300	10.4800
	Mean $\|r\|$	0.95071	0.95473	0.9618	0.9317	0.9689	0.9762
	SD $\|r\|$	0.01593	0.01406	0.0147	0.0216	0.0104	0.0098
	MMSE	0.00070	0.00070	0.0006	0.0008	0.0004	0.0004

Table 8. Simulation results for the ALSIQR, LSIQR, SIQR, ALLQR, LLQR and LQR based on the linear model in example 3 model 6, $\sigma = 1$ and $\sigma = 3$

		LQR	LLQR	ALLQR	SIQR	LSIQR	ALSIQR
$\tau = 0.10$							
$\sigma = 1$	Ave 0's	2.0300	7.4800	10.5500	2.6500	9.6800	11.0000
	Mean $\|r\|$	0.9944	0.9948	0.9959	0.9776	0.9967	0.9963
	SD $\|r\|$	0.0015	0.0016	0.0018	0.0099	0.0012	0.0014
	MMSE	0.0079	0.0075	0.0074	0.0028	0.0027	0.0016
$\sigma = 3$	Ave 0's	6.7500	6.9500	8.7500	9.8500	9.4300	11.0300
	Mean $\|r\|$	0.9519	0.9544	0.9581	0.9482	0.9799	0.9824
	SD $\|r\|$	0.0179	0.0171	0.0177	0.0165	0.0097	0.0096
	MMSE	0.0264	0.0239	0.0220	0.0154	0.0136	0.0147
$\tau = 0.50$							
$\sigma = 1$	Ave 0's	2.1300	7.4000	10.8300	2.9000	10.1000	10.1000
	Mean $\|r\|$	0.9961	0.9963	0.9962	0.9758	0.9965	0.9963
	SD $\|r\|$	0.0011	0.0011	0.0017	0.0051	0.0011	0.0016
	MMSE	0.0006	0.0005	0.0005	0.0005	0.0004	0.0004
$\sigma = 3$	Ave 0's	6.7800	7.1500	9.5000	10.9000	10.5000	12.3500
	Mean $\|r\|$	0.9668	0.9683	0.9720	0.9501	0.9801	0.9821
	SD $\|r\|$	0.0111	0.0113	0.0104	0.0169	0.0070	0.0084
	MMSE	0.0076	0.0072	0.0059	0.0064	0.0058	0.0051
$\tau = 0.50$							
	Ave 0's	2.1300	7.7300	11.0000	2.4000	9.2500	11.4000
$\sigma = 1$	Mean $\|r\|$	0.9965	0.9967	0.9966	0.9812	0.9969	0.9967
	SD $\|r\|$	0.0014	0.0014	0.0011	0.0064	0.0010	0.0010
	MMSE	0.0002	0.0002	0.0002	0.0003	0.0001	0.0001
$\sigma = 3$	Ave 0's	6.9000	7.2800	9.7300	10.4800	10.8000	12.7300
	Mean $\|r\|$	0.9740	0.9753	0.9791	0.9548	0.9812	0.9819
	SD $\|r\|$	0.0083	0.0082	0.0081	0.0174	0.0054	0.0073
	MMSE	0.0021	0.0020	0.0022	0.0041	0.0018	0.0017

Table 9. Simulation results for the ALSIQR, LSIQR, SIQR and NQR based on the nonlinear model in example 4

	NQR	SIQR	LSIQR	ALSIQR
$\tau = 0.10$				
Ave 0's	1.0900	1.0600	3.1000	3.2400
Mean $\|r\|$	0.8541	0.9951	0.9987	0.9991
SD $\|r\|$	0.0717	0.0036	0.0010	0.0007
MMSE	0.0059	0.0003	0.0001	0.0001
$\tau = 0.25$				
Ave 0's	1.0700	1.0300	3.1400	3.2500
Mean $\|r\|$	0.8352	0.9946	0.9985	0.9991
SD $\|r\|$	0.0788	0.0041	0.0011	0.0009
MMSE	0.0030	0.0003	0.0002	0.0001
$\tau = 0.50$				
Ave 0's	1.1100	1.0800	3.2500	3.3300
Mean $\|r\|$	0.8235	0.9948	0.9985	0.9990
SD $\|r\|$	0.0793	0.0036	0.0010	0.0008
MMSE	0.0065	0.0003	0.0002	0.0001

Table 10. The results of MAD and the SD of the prediction errors for estimated quantiles which are estimated by the ALSIQR, LSIQR, SIQR, ALLQR, LLQR and LQR based on air pollution data for τ = (0.10,0.25,0.50)

		MAD	SD of the prediction errors
$\tau = 0.50$	ALSIQR	0.6532	1.0170
	LSIQR	0.6672	1.0105
	SIQR	0.6645	1.0154
	ALLQR	1.2861	1.0297
	LLQR	1.2374	1.0471
	LQR	1.2372	1.0475
$\tau = 0.50$	ALSIQR	0.6443	0.9852
	LSIQR	0.6635	0.9750
	SIQR	0.6729	0.9751
	ALLQR	0.7200	1.0112
	LLQR	0.7239	1.0332
	LQR	0.7240	1.0095
$\tau = 0.50$	ALSIQR	0.6464	1.0019
	LSIQR	0.9466	1.0484
	SIQR	1.2971	1.0699
	ALLQR	0.5846	0.9562
	LLQR	0.5869	0.9527
	LQR	0.5869	0.9527

LSIQR and ALSIQR as compared to NQR and SIQR give better results according to MMSE and SD for all values of τ, as Table 9 results show.

In general, the ALSIQR, LSIQR and SIQR produce precise estimates and they are more significantly efficient than the other methods.

It can be observed that the ALSIQR, LSIQR and SIQR give a lower MMSE and bigger $|r_i|$ than the other methods. The variations in estimates of the proposed methods are approximately same in the most of cases and less than the variations in the estimate of the other methods. Most noticeably, when $\tau = 0.10$ and $\tau = 0.25$, the ALSIQR, LSIQR and SIQR are more considerably efficient than the other methods. From Table 3-9, in term of variables selection and according to Ave 0's, it is obvious that the Ave 0's for the ALSIQR and LSIQR methods is close to the true number.

Air Pollution (A.R) Data

In this section, the ALSIQR, LSIQR and SIQR have been illustrated through an analysis of A.R data. The A.R data is online at the website http://lib.stat.cmu.edu/datasets/NO2.dat. The response Y is hourly values of LOG of the concentration of NO2. The $p = 7$ predictors X are LOG of the number of cars/hour (x_1), temperature 2 m above ground (x_2), wind speed (x_3), the difference in temperature between 2 and 25 m above ground (x_4), wind direction (x_5), hour of day (x_6) and day number (x_7). The predictors and the response have been standardized.

Table 10 reports the results of MAD (median absolute difference between $\hat{\beta}^T X$ and y) and SD of the prediction errors for estimated quantiles by all of the studied methods based on A.R for $\tau = (0.10, 0.25, 0.50)$. It is clear that the ALSIQR, LSIQR and SIQR have less MAD and SD of the prediction errors than the other methods especially for τ = 0.0.10 and 0.25. This confirms the statement that the proposed methods do well in the extreme quantiles. The results of the numerical examples and the A.R analysis suggests that the ALSIQR, LSIQR and SIQR perform well.

Conclusion

The current study proposes three methods, SIQR, LSIQR and ALSIQR. The ALSIQR, LSIQR and SIQR have been compared with ALLQR, LLQR, LQR and NQR under different situations. In order to examine the performance of the SIQR, LSIQR and ALSIQR, numerical examples were conducted based on the models as described in section 5. It has been concluded based on the simulation studies and A.R data, that the SIQR, LSIQR and ALSIQR more advantageous in comparison to ALLQR, LLQR, LQR and NQR and thus the authors believe that the SIQR, LSIQR and ALSIQR are useful practically.

Acknowledgement

Our thanks are going to the anonyms referees for the helpful comments.

Funding Information

This article is supported by University of Al-Qadisiyah, Iraq.

Author's Contributions

The manuscript has been done according to the corporation between the authors in each point.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

Alkenani, A. and K. Yu, 2013. Penalized single-index quantile regression. Int. J. Stat. Probab., 2: 12-30. DOI: 10.5539/ijsp.v2n3p12

Aragon, Y. and J. Saracco, 1997. Sliced Inverse Regression (SIR): An appraisal of small sample alternatives to slicing. Comput. Stat., 12: 109-130.

Chaudhuri, P., 1991. Global nonparametric estimation of conditional quantile functions and their derivative. J. Multivariate Analysis, 39: 246-269. DOI: 10.1016/0047-259X(91)90100-G

Cook, R.D., 1996. Graphics for regressions with a binary response. J. Am. Stat. Assoc., 91: 983-992. DOI: 10.1080/01621459.1996.10476968

Cook, R., 1998. Regression Graphics: Ideas for Studying Regressions Through Graphics. 1st Edn., John Wiley and Sons, New York, ISBN-10: 0471193658, pp: 349.

Cook, R.D., 2004. Testing predictor contributions in sufficient dimension reduction. Ann. Stat., 32: 1061-1092. DOI: 10.1214/009053604000000292

Dette, H. and R. Scheder, 2011. Estimation of additive quantile regression. Ann. Insti. Stat. Math., 63: 245-265. DOI: 10.1007/s10463-009-0225-5

Fan, J. and R.Z. Li, 2001. Variable selection via nonconcave penalized likelihood and its oracle properties. J. Am. Stat. Assoc., 96: 1348-1360. DOI: 10.1198/016214501753382273

Efron, B., T. Hastie, I. Johnstone and R. Tibshirani, 2004. Least angle regression. Ann. Stat., 32: 407-499. DOI: 10.1214/009053604000000067

Gannoun, A., S. Girard and J. Saracco, 2004. Sliced inverse regression in reference curves estimation. Comput. Stat. Data Analysis, 46: 103-122. DOI: 10.1016/S0167-9473(03)00141-5

Horowitz, J.L and S. Lee, 2005. Nonparametric estimation of an additive quantile regression model. J. Am. Stat. Assoc., 100: 1238-1249. DOI: 10.1198/016214505000000583

Li, K., 1991. Sliced inverse regression for dimension reduction. J. Am. Stat. Assoc., 86: 316-342. DOI: 10.1080/01621459.1991.10475035

Li, L., 2007. Sparse sufficient dimension reduction. Biometrika, 94: 603-613. DOI: 10.1093/biomet/asm044

Li, L. and C.J. Nachtsheim, 2006. Sparse sliced inverse regression. Technometrics, 48: 503-510. DOI: 10.1198/004017006000000129

Li, L. and X. Yin, 2008. Sliced inverse regression with regularizations. Biometrics, 64: 124-131. DOI: 10.1111/j.1541-0420.2007.00836.x

Li, Y. and J. Zhu, 2008. L_1-Norm quantile regression. J. Comput. Graphical Stat., 17: 163-185. DOI: 10.1198/106186008X289155

Ni, L., R.D. Cook and C.L. Tsai, 2005. A note on shrinkage sliced inverse regression. Biometrika, 92: 242-247. DOI: 10.1093/biomet/92.1.242

Tibshirani, R., 1996. Regression shrinkage and selection via the Lasso. J. Royal Stat. Society Ser. B, 58: 267-288.

Wu, T.Z., K. Yu and Y. Yu, 2010. Single-index quantile regression. J. Multivariate Analysis, 101: 1607-1621. DOI: 10.1016/j.jmva.2010.02.003

Wu, Y. and Y. Liu, 2009. Variable selection in quantile regression. Stat. Sinica, 19: 801-817.

Yebin, C., J.G.D. Gooijer and D. Zerom, 2011. Efficient estimation of an additive quantile regression model. Scandinavian J. Stat., 38: 46-62. DOI: 10.1111/j.1467-9469.2010.00706.x

Yu, K., Z. Lu and J. Stander, 2003. Quantile regression: Applications and current research areas. Statistician, 52: 331-350. DOI: 10.1111/1467-9884.00363

Zhu, L., B. Miao and H. Peng, 2006. On sliced inverse regression with high-dimensional covariates. J. Am. Stat. Assoc., 101: 630-643. DOI: 10.1198/016214505000001285

Zou, H., 2006. The adaptive lasso and its oracle properties. J. Am. Stat. Assoc., 101: 1418-1429. DOI: 10.1198/016214506000000735