A $\nu$- support vector quantile regression model with automatic accuracy control

Pritam Anand · Reshma Rastogi (nee Khemchandani) · Suresh Chandra.

Received: date / Accepted: date

Abstract  This paper proposes a novel $\nu$-support vector quantile regression ($\nu$-SVQR) model for the quantile estimation. It can facilitate the automatic control over accuracy by creating a suitable asymmetric $\epsilon$-insensitive zone according to the variance present in data. The proposed $\nu$-SVQR model uses the $\nu$ fraction of training data points for the estimation of the quantiles. In the $\nu$-SVQR model, training points asymptotically appear above and below of the asymmetric $\epsilon$-insensitive tube in the ratio of $1 - \tau$ and $\tau$. Further, there are other interesting properties of the proposed $\nu$-SVQR model, which we have briefly described in this paper. These properties have been empirically verified using the artificial and real world dataset also.

Keywords  Quantile Regression, pinball loss function, Support Vector Machine, $\epsilon$-insensitive loss function.

Pritam Anand
Research Scholar,
Faculty of Mathematics and Computer Science,
South Asian University, New Delhi-110021.
E-mail: ltpritamanand@gmail.com.

R. Rastogi (nee Khemchandani)
Assistant Professor,
Faculty of Mathematics and Computer Science,
South Asian University, New Delhi-110021.
E-mail: reshma.khemchandani@sau.ac.in.
Tel No. 01124195147

Suresh Chandra
Ex-Faculty, Department of Mathematics
Indian Institute of Technology Delhi, New Delhi-110016.
E-mail: chandras@maths.iitd.ac.in
1 Introduction

Given the training set $T = \{(x_i, y_i) : x_i \in \mathbb{R}^n, y_i \in \mathbb{R}, i = 1, 2, \ldots, l \}$ and $\tau \in [0, 1]$, the problem of quantile regression is to estimate a real valued function $f_\tau(x)$ such that a proportion $\tau$ of $y/x$ will be lying below of the estimate $f_\tau(x)$. For $\tau = 0.5$, the problem is equivalent to median estimation. The estimation of $f_\tau(x)$ is difficult but, more informative than estimation of only mean regression $f(x)$. The estimation of $f_\tau(x)$ for different values of $\tau$ can briefly describe the different characteristics of the conditional distribution of $y/x$. In many real world problems, the estimation of mean regression $f(x)$ is not required or enough, rather they require the estimation of quantile $f_\tau(x)$.

The study of quantile regression problem has initially been started in 1978 by Koenkar and Bassett[1]. Later, it has been briefly discussed and described by Koenker in his book (Koenker, [2]). Koenkar and Bassett [1] proposed the pinball loss function for the estimation of the quantile function $f_\tau(x)$. For a given quantile $\tau \in (0, 1)$, the pinball loss function was an asymmetric loss function suitable for quantile estimation. It was given by

$$P_\tau(u) = \begin{cases} 
\tau u & \text{if } u \geq 0, \\
(\tau - 1)u & \text{otherwise.}
\end{cases} \tag{1}$$

Support Vector Regression (SVR) models (Vapnik et al.,[7])(Drucker et al.,[8]),(Gunn, [14]) are one of the most popular regression model which can estimate the mean regression function $f(x)$ efficiently. SVR model commonly solves a Convex Program which guarantees the global optimal solution. These models have been widely used in solving real world problems of diverse domain.

Takeuchi et al [6] initiated the study of the quantile regression problem in a non-parametric framework on the line of SVR models. They have proposed Support Vector Quantile Regression (SVQR) model in which they have minimized the pinball loss function in SVR type optimization problem for estimation of the quantile function $f_\tau(x)$. The obtained solution of SVQR model is not sparse as every training data points are allowed to contribute in the empirical risk which is measured by the asymmetric pinball loss function.

Researchers have attempted to extend the SVQR model on the line of $\epsilon$-SVR model for increasing its generalization ability as well as obtaining the sparse solution. For this, they have attempted to propose the $\epsilon$-insensitive pinball loss functions to incorporate the concept of $\epsilon$-insensitive zone in the asymmetric pinball loss function.

At first, Takeuchi and Furuhashi considered the $\epsilon$-insensitive pinball loss function for estimation of the non-crossing quantile in their work (Takeuchi and Furuhashi, [10]). Further, Hu et al, had also considered the similar kind of $\epsilon$-insensitive pinball loss function in their work (Hu et al, [11]) for estimation of quantiles. However, the $\epsilon$-insensitive zone in these pinball loss function was symmetric. The use of the symmetric $\epsilon$-insensitive zone in the asymmetric pinball loss function failed to perform well for estimation of quantiles.
Soek et al. have first considered the asymmetric $\epsilon$-insensitive zone in the pinball loss function in their proposed $\epsilon$-sensitive pinball loss function (Soek et al., [12]). Later on, Park and Kim [13] has also proposed a similar kind of loss function in their work (Park and Kim, [13]). But problem of these pinball loss function was that they failed to provide a suitable $\epsilon$-insensitive zone for every value of $\tau$.

Anand et al. have proposed an asymmetric $\epsilon$-insensitive pinball loss function in their work (Anand et al., [15]) which extends the concept of $\epsilon$-insensitive zone in the pinball loss function in true sense. The asymmetric $\epsilon$-insensitive pinball loss can obtain a suitable $\epsilon$-insensitive zone of fixed width for every values of $\tau$. The $\epsilon$ -insensitive zone was partitioned using $\tau$ value in the asymmetric $\epsilon$-insensitive pinball loss function. Using the asymmetric $\epsilon$-insensitive pinball loss function, they have proposed $\epsilon$-SVQR model which can obtain better generalization ability than existing SVQR models and successfully brings the sparsity back in the SVQR model.

However, the $\epsilon$-SVQR model (Anand et al., [15]) requires a good choice of value of $\epsilon$ for obtaining the better prediction of quantiles. A bad choice of $\epsilon$ can distort the performance of the $\epsilon$-SVQR model (Anand et al., [15]).

This paper proposes an efficient SVQR model which appropriately trade-off the total width of the asymmetric $\epsilon$-insensitive zone in its optimization problem via the user defined parameter $\nu$. The proposed model has been termed with $\nu$-Support Vector Quantile Regression ($\nu$-SVQR) model. The $\nu$-SVQR model can adjust the overall width of asymmetric $\epsilon$-insensitive zone such that at most $\nu$ fraction of training data points lie outside of it. This capability of $\nu$ -SVQR enables it to automatically adjust the width of the $\epsilon$-insensitive zone according to the variance present in the data without adjusting any parameter. In the $\nu$-SVQR model, training points asymptotically appear above and below of the asymmetric $\epsilon$-insensitive tube in the ratio of $1 - \tau$ and $\tau$. Further, there are other interesting asymptotic properties of $\nu$-SVQR model which we have briefly described in this paper. Several experiments on artificial as well as UCI datasets have been performed to empirically verify claims made in this paper.

The rest of this paper is organized as follows. Section-2 briefly describes the standard Support Vector Quantile Regression(SVQR) model[6] and $\epsilon$-Support Vector Quantile Regression ($\epsilon$-SVQR) model[15]. In Section-3, we present our proposed $\nu$-Support Vector Quantile Regression ($\nu$-SVQR) model and its different properties. Section-4 contains the numerical results obtained by different nature of experiments carried on artificial as well as real world datasets to empirically verify the properties of proposed $\nu$-SVQR model.

2 Support Vector Quantile Regression models

For the training set $T = \{(x_i, y_i) : x_i \in \mathbb{R}^n, y_i \in \mathbb{R}, i = 1, 2...l \}$ and the quantile $\tau \in (0, 1)$ , the SVQR model estimates the function $f_\tau(x) = w^T \phi(x) + b$ in the feature space for the estimation of the $\tau$th quantile, where
\( \phi : \mathbb{R}^n \to \mathcal{H} \) is a mapping from the input space to a higher dimensional feature space \( \mathcal{H} \).

### 2.1 Standard Support Vector Quantile Regression model

The standard Support Vector Quantile regression model minimizes

\[
\min_{w,b} \frac{1}{2} ||w||^2 + C \sum_{i=1}^{l} L_\tau(y_i - (w^T x_i + b)),
\]

where \( L_\tau(v) \) is the asymmetric pinball loss function which is given by

\[
L_\tau(v) = \begin{cases} 
\tau v & \text{if } v > 0, \\
(\tau - 1)v & \text{otherwise}.
\end{cases}
\]

Using the \( l \)-dimensional variables \( \xi = (\xi_1, \xi_2, ..., \xi_l) \) and \( \xi^* = (\xi^*_1, \xi^*_2, ..., \xi^*_l) \), the optimization problem (2) can be equivalently converted to following Quadratic Programming Problem (QPP)

\[
\min_{(w,b,\xi,\xi^*)} \frac{1}{2} ||w||^2 + C \sum_{i=1}^{l} (\tau \xi_i + (1 - \tau)\xi^*_i)
\]

subject to,

\[
y_i - (w^T \phi(x_i) + b) \leq \xi_i,
\]

\[
(w^T \phi(x_i) + b) - y_i \leq \xi^*_i,
\]

\[
\xi_i \geq 0, \quad \xi^*_i \geq 0, \quad i = 1, 2, ..., l.
\]

Here \( C \geq 0 \) is a user defined parameter which is used to find a good trade-off between empirical risk and model complexity of estimator. The QPP (4) of standard SVQR model can be easily solved by solving its corresponding Wolfe dual problem. More detail about standard SVQR model can be found in (Takeuchi et al.,[6]).

### 2.2 \( \epsilon \)- Support Vector Quantile Regression model

Anand et al.[15] have proposed an asymmetric \( \epsilon \)-insensitive pinball loss function which can obtain a suitable asymmetric \( \epsilon \)-insensitive zone for every values of \( \tau \). The asymmetric \( \epsilon \)-insensitive pinball loss function is given by

\[
L_\tau^\epsilon(u) = \max(- (1 - \tau)(u + \tau \epsilon), \quad 0, \quad \tau(u - (1 - \tau)\epsilon)).
\]

It can be better understood in the following form

\[
L_\tau^\epsilon(y_i, x_i, w, b) = \begin{cases} 
- (1 - \tau)(y_i - (w^T x_i + b) + \tau \epsilon), & \text{if } y_i - (w^T x_i + b) < -\tau \epsilon. \\
0, & \text{if } -\tau \epsilon \leq y_i - (w^T x_i + b) \leq (1 - \tau)\epsilon. \\
\tau(y_i - (w^T x_i + b) - (1 - \tau)\epsilon), & \text{if } y_i - (w^T x_i + b) > (1 - \tau)\epsilon.
\end{cases}
\]
Fig. 1: The asymmetric $\epsilon$-pinball loss function for (a) $\tau = 0.1$ (b) $\tau = 0.2$ (c) $\tau = 0.5$ and (d) $\tau = 0.8$ with fixed $\epsilon = 1$.

Figure (1) shows that the asymmetric $\epsilon$-insensitive pinball loss function can generate the suitable asymmetric $\epsilon$-insensitive zone for different values of $\tau$.

The $\epsilon$-SVQR model minimizes

$$
\min_{(w,b)} \frac{1}{2} ||w||^2 + C \sum_{i=1}^{l} L_{\tau}^\epsilon(y_i, x_i, w, b)
$$

which can be equivalently converted to following QPP

$$
\min_{(w,b,\xi,\xi^*)} \frac{1}{2} ||w||^2 + C \sum_{i=1}^{l} (\tau \xi_i + (1 - \tau) \xi_i^*)
$$

subject to,

$$
\begin{align*}
y_i - (w^T \phi(x_i) + b) & \leq (1 - \tau) \epsilon + \xi_i, \\
(w^T \phi(x_i) + b) - y_i & \leq \tau \epsilon + \xi_i^*, \\
\xi_i & \geq 0, \quad \xi_i^* \geq 0, \quad i = 1, 2, \ldots, l.
\end{align*}
$$

In the $\epsilon$-SVQR model, $\epsilon \geq 0$ is the user defined parameter and a good value of $\epsilon$ is required beforehand for the efficient estimate of quantiles. For the solution
of the $\epsilon$-SVQR primal problem (8), we obtain its corresponding Wolfe dual problem as follows

$$\min_{\alpha, \beta} \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} (\alpha_i - \beta_j)K(x_i, x_j)(\alpha_j - \beta_i) - \sum_{i=1}^{l} (\alpha_i - \beta_i)y_i + \sum_{i=1}^{l} ((1 - \tau)\alpha_i + \tau \epsilon \beta_i)$$

subject to,

$$\sum_{i=1}^{l} (\alpha_i - \beta_i) = 0,$$

$$0 \leq \alpha_i \leq C \tau, \quad i = 1, 2, ... l,$$

$$0 \leq \beta_i \leq C (1 - \tau), \quad i = 1, 2, ... l.$$  \hspace{1cm} (9)

After obtaining the solution of the dual problem (9), we can estimate $f_\tau(x)$, for any test data point $x \in \mathbb{R}^n$ using

$$f_\tau(x) = \sum_{i=1}^{l} (\alpha_i - \beta_i)K(x, x_i) + b.$$  \hspace{1cm} (10)

3 Proposed $\nu$-Support Vector Quantile Regression model

The proposed $\nu$-SVQR model minimizes

$$\min_{(w, b, \epsilon, \xi, \xi^*)} \frac{1}{2} ||w||^2 + C (\nu \tau (1 - \tau) \epsilon + \frac{1}{l} \sum_{i=1}^{l} (\tau \xi_i + (1 - \tau) \xi_i^*))$$

subject to,

$$y_i - (w^T \phi(x_i) + b) \leq (1 - \tau) \epsilon + \xi_i,$$

$$(w^T \phi(x_i) + b) - y_i \leq \tau \epsilon + \xi^*_i,$$

$$\xi_i \geq 0, \quad \xi^*_i \geq 0, \quad \epsilon \geq 0, \quad i = 1, 2, ... l.$$  \hspace{1cm} (11)

where $C \geq 0$ and $\nu \geq 0$ are user defined parameters.

For solving the primal problem (11) efficiently, we need to derive its Wolfe dual problem. The Lagrangian function for the primal problem (11) is obtained as

$$L(w, b, \epsilon, \xi, \xi^*, \alpha_i, \beta_i, \gamma_i, \lambda_i, \eta) = \frac{1}{2} ||w||^2 + C (\nu \tau (1 - \tau) \epsilon + \frac{1}{l} \sum_{i=1}^{l} (\tau \xi_i + (1 - \tau) \xi_i^*))$$

$$+ \sum_{i=1}^{l} \alpha_i (y_i - (w^T \phi(x_i) + b) - (1 - \tau) \epsilon - \xi_i) + \sum_{i=1}^{l} \beta_i ((w^T \phi(x_i) + b) - y_i - \tau \epsilon - \xi^*_i)$$

$$- \sum_{i=1}^{l} \gamma_i \xi_i - \sum_{i=1}^{l} \lambda_i \xi^*_i - \eta \epsilon$$  \hspace{1cm} (12)
We can now note the KKT conditions for (11) as follows

$$\frac{\partial L}{\partial w} = w + \sum_{i=1}^{l} (\beta_i - \alpha_i) \phi(x_i) = 0 \implies w = \sum_{i=1}^{l} (\alpha_i - \beta_i) \phi(x_i)$$

(13)

$$\frac{\partial L}{\partial \alpha_i} = \sum_{i=1}^{l} (\beta_i - \alpha_i) = 0.$$ 

(14)

$$\frac{\partial L}{\partial \beta_i} = \frac{\gamma_i}{T} \tau - \alpha_i - \gamma_i = 0, \; i = 1, 2, ..., l.\quad (15)$$

$$\frac{\partial L}{\partial \gamma_i} = \frac{\gamma_i}{T} (1 - \tau) - \beta_i - \lambda_i = 0, \; i = 1, 2, ..., l.\quad (16)$$

$$\frac{\partial L}{\partial \beta_i} = C \nu \tau (1 - \tau) - (1 - \tau) \sum_{i=1}^{l} \alpha_i - \tau \sum_{i=1}^{l} \beta_i - \eta = 0 \quad (17)$$

$$\alpha_i (y_i - (w^T \phi(x_i) + b) - (1 - \tau) \epsilon - \xi_i) = 0, \; i = 1, 2, ..., l.\quad (18)$$

$$\beta_i ((w^T \phi(x_i) + b) - y_i - \tau \epsilon - \xi_*^i) = 0, \; i = 1, 2, ..., l.\quad (19)$$

$$\gamma_i \xi_i = 0, \; \lambda_i \xi_*^i = 0 \; i = 1, 2, ..., l.\quad (20)$$

$$\eta \epsilon = 0.\quad (21)$$

$$y_i - (w^T \phi(x_i) + b) \leq (1 - \tau) \epsilon + \xi_i, \; i = 1, 2, ..., l.\quad (22)$$

$$(w^T \phi(x_i) + b) - y_i \leq \tau \epsilon + \xi_*^i, \; i = 1, 2, ..., l.\quad (23)$$

$$\epsilon \geq 0, \; \xi_i \geq 0, \; \xi_*^i \geq 0, \; i = 1, 2, ..., l.\quad (24)$$

Making the use the above KKT conditions, the Wolfe dual problem of the primal problem (11) can be obtained as follows

$$\min_{\alpha, \beta} \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} (\alpha_i - \beta_j) K(x_i, x_j) (\alpha_j - \beta_i) - \sum_{i=1}^{l} (\alpha_i - \beta_i) y_i$$

subject to,

$$\sum_{i=1}^{l} (\alpha_i - \beta_i) = 0,$$

$$(1 - \tau) \sum_{i=1}^{l} \alpha_i + \tau \sum_{i=1}^{l} \beta_i \leq C \nu \tau (1 - \tau),$$

$$0 \leq \alpha_i \leq \frac{\gamma_i}{T} \tau, \; i = 1, 2, ..., l,$$

$$0 \leq \beta_i \leq \frac{\gamma_i}{T} (1 - \tau), \; i = 1, 2, ..., l.\quad (25)$$

The KKT conditions (13)- (24) will help us to discover the various characteristics of the proposed $\epsilon$-SVQR model. At first, we shall state following preposition.

**Preposition 1.** $\alpha_i \beta_i = 0$ and $\xi_i \xi_*^i = 0$ holds $\forall i=1,2,...,l.$

Proof:- If possible, let us suppose there exists an index $i$ such that $\alpha_i \beta_i \neq 0$ holds. It implies that $\alpha_i \neq 0$ and $\beta_i \neq 0$. Therefore, from the KKT condition (18) and (19) we can obtain

$$(y_i - (w^T \phi(x_i) + b) - (1 - \tau) \epsilon - \xi_i) = 0 \quad (26)$$

and

$$((w^T \phi(x_i) + b) - y_i - \tau \epsilon - \xi_*^i) = 0.\quad (27)$$

Adding equation (26) and (27) gives $\xi_*^i + \xi_i = -\epsilon$ which is possible only when either $\xi_i < 0$ or $\xi_*^i < 0$. But, the KKT condition (24) requires $\xi_i \geq 0, \; \xi_*^i \geq$
0, for \( i = 1, 2, ..., l \) which contradicts our assumption. This proves \( \alpha_i \beta_i = 0 \)
\( \forall i=1,2,...l \).

On the similar line, let us suppose that there exists an index \( i \) for which \( \xi_i \xi_i^* \neq 0 \). It means that \( \xi_i \neq 0 \) and \( \xi_i^* \neq 0 \) for which we can obtain \( \gamma_i = 0 \) and \( \lambda_i = 0 \) from KKT condition (20). For \( \gamma_i = 0 \) and \( \lambda_i = 0 \), we will obtain \( \alpha_i = \xi \tau \) and \( \beta_i = \xi (1-\tau) \) from the KKT conditions (15) and (16) respectively, which is not possible as we have already proven that \( \alpha_i \beta_i = 0 \) \( \forall i=1,2,...l \). This proves \( \xi_i \xi_i^* = 0 \) \( \forall i=1,2,...l \).

**Proposition 2.** For all those data points \((x_i, y_i)\), which lie inside or boundary of the asymmetric \( \epsilon \)-insensitive tube, the corresponding \( \xi_i \) and \( \xi_i^* \) will take zero value.

**Proof:** The data point \((x_i, y_i)\) lying inside or boundary of the asymmetric \( \epsilon \)-insensitive tube must satisfy
\[
y_i - (w^T \phi(x_i) + b) - (1-\tau)\epsilon \leq 0
\]
and
\[
(w^T \phi(x_i) + b) - y_i - \tau \epsilon \leq 0
\]

If possible, let us suppose that \( \xi_i \neq 0 \) which means that \( \xi_i > 0 \) (as the KKT condition (24) requires \( \xi \geq 0 \)). Since \( \xi_i > 0 \), we can obtain \( \gamma_i = 0 \) and further \( \alpha_i > 0 \) by using the KKT conditions (20) and (18) respectively. For \( \alpha_i > 0 \), the KKT condition (18) implies that
\[
y_i - (w^T \phi(x_i) + b) - (1-\tau)\epsilon = \xi_i,
\]
which is not possible as \( \xi_i > 0 \).

On the similar line, we can show that \( \xi_i^* \) also cannot take non-zero values.

It is also easy to prove that data point, which lie outside of the asymmetric \( \epsilon \)-insensitive tube, the corresponding \( \xi_i \) or \( \xi_i^* \) will take positive value.

**Proposition 3.** For the data point \((x_i, y_i)\), which lie inside of the asymmetric \( \epsilon \)-insensitive tube, the corresponding \( \alpha_i \) and \( \beta_i \) will take zero value.

**Proof:** The data point \((x_i, y_i)\) lying inside of the \( \epsilon \)-tube, the \( \xi_i \) and \( \xi_i^* = 0 \) which means
\[
y_i - (w^T \phi(x_i) + b) - (1-\tau)\epsilon - \xi_i \leq 0,
\]
and
\[
(w^T \phi(x_i) + b) - y_i - \tau \epsilon - \xi_i^* \leq 0.
\]
For which the use of the KKT condition (18) and (19) will let us obtain \( \alpha_i \) and \( \beta_i = 0 \).

**Proposition 4.** For the data point \((x_i, y_i)\), lying above of the \( \epsilon \)-tube, \( \alpha_i = \xi \tau \) and \( \beta_i = 0 \). For the data point \((x_i, y_i)\), lying below of the \( \epsilon \)-tube, \( \alpha_i = 0 \) and \( \beta_i = \xi (1-\tau) \).

**Proof:** The data point \((x_i, y_i)\), lying above of the \( \epsilon \)-tube will hold
\[
y_i - (w^T \phi(x_i) + b) - (1-\tau)\epsilon > 0,
\]
for which the corresponding \( \xi_i \) will take positive value for satisfying the KKT condition (22). For \( \xi_i > 0 \), we can get \( \gamma_i = 0 \) from (20) and further can obtain
\[ \alpha_i = \frac{C}{T} \tau \] from the KKT condition[15]. Further \( \beta_i \) will take zero value as \( \alpha_i \beta_i = 0 \).

On the similar line, we can prove that the data point \((x_i, y_i)\), lying below of the \( \epsilon \)-tube, \( \alpha_i = 0 \) and \( \beta_i = \frac{C}{T}(1 - \tau) \).

**Preposition 5.** For \( 0 < \alpha_i < \frac{C}{T} \tau \) \( (0 < \beta_i < \frac{C}{T}(1 - \tau)) \), the corresponding data point \((x_i, y_i)\) will be lying on the upper (lower) boundary of the asymmetric \( \epsilon \)-insensitive tube.

**Proof:** For \( 0 < \alpha_i < \frac{C}{T} \tau \), the \( \gamma_i > 0 \) from KKT Condition [15] which implies \( \xi_i = 0 \). Further, for \( \alpha_i > 0 \), we can obtain

\[ y_i - (w^T \phi(x_i) + b) = (1 - \tau) \epsilon, \] (34)

which means that data point \((x_i, y_i)\) will be lying on the upper boundary of the asymmetric \( \epsilon \)-insensitive tube. On the similar line, we can obtain that for \( 0 < \beta_i < \frac{C}{T}(1 - \tau) \), the corresponding data point \((x_i, y_i)\) will be lying on the below boundary of the asymmetric \( \epsilon \)-insensitive tube.

Now, we can argue that the data point \((x_i, y_i)\), which are lying outside of the \( \epsilon \)-tube

**Remark 1.** For \( \epsilon > 0 \), the \( \eta \) will take zero value and the inequalities constraint \( (1 - \tau) \sum_{i=1}^{l} \alpha_i + \tau \sum_{i=1}^{l} \beta_i \leq C \nu \tau (1 - \tau) \) of the dual problem [25] will get converted to the equaltiy constraint

\[ (1 - \tau) \sum_{i=1}^{l} \alpha_i + \tau \sum_{i=1}^{l} \beta_i = C \nu \tau (1 - \tau). \] (35)

Now, we shall term the data points which are lying outside of the asymmetric \( \epsilon \)-tube with ‘Errors’. The data points which are lying outside of the asymmetric \( \epsilon \)-tube as well as boundary of the tube is termed with the ‘support vectors’. These data points only contributes for the construction of the final regressor.

**Preposition 6.** Suppose the \( \nu \)-SVQR is applied to some dataset and resulting \( \epsilon \) is non zero then the following statements hold.

(a) \( \nu \) is upper bound on the fraction of of Errors.

(b) \( \nu \) is lower bound on the fraction of of Support vectors.

**Proof:**- Let us suppose that there are \( m_1 \) and \( m_2 \) data points which are lying above and below of the asymmetric \( \epsilon \)-tube respectively. For the data point, lying above of the asymmetric \( \epsilon \)-tube, only \( \alpha_i \) will take the value \( \frac{C}{T} \tau \). For the data point, lying below of the asymmetric \( \epsilon \)-tube, only \( \beta_i \) will take the value \( \frac{C}{T}(1 - \tau) \). The data point \((x_i, y_i)\) lying on the upper (lower) boundary of the asymmetric \( \epsilon \)-insensitive tube will be taking \( 0 < \alpha_i < \frac{C}{T} \tau \) \( (0 < \beta_i < \frac{C}{T}(1 - \tau)) \) values.

For \( \epsilon > 0 \), we can obtain from (35)

\[ m_1 \frac{C}{T} (1 - \tau) \tau + m_2 \frac{C}{T} (1 - \tau) \tau \leq C \nu \tau (1 - \tau) \]

which implies that \( \frac{1}{T}(m_1 + m_2) \leq \nu. \)
Furthermore, there should exist at least $m_1$ and $m_2$ data points lying above and below of the asymmetric $\epsilon$-tube which would satisfy the equality. For these data points we have,

$$m_1 \frac{C}{l} (1 - \tau) \tau + m_2 \frac{C}{l} (1 - \tau) \tau = C \nu \tau (1 - \tau)$$

which implies that $\frac{1}{l}(m_1 + m_2) = \nu$. It further means that there should at least $\nu$ fraction of the support vectors.

**Remark 2.** Asymptotically, the $\nu$ equals the fraction of support vectors and errors. The probability of the data point lying on the boundary of the asymmetric $\epsilon$-tube becomes zero asymptotically. This statement can be proved under certain condition similar to the proof of the Prepostion 1 (iii) given in (Scholkopf, [16]). But however, in this paper we shall empirically verify that the $\nu$ equals the fraction of support vectors and errors asymptotically.

**Remark 3.** Asymptotically, the data points appear above and below of the asymmetric $\epsilon$-tube in the ratio of $1 - \tau$ and $\tau$ respectively in the $\nu$-SVQR model. It means that for the large value of $l$, there would be $l \nu (1 - \tau)$ and $l \nu \tau$ data points lying above and below of the asymmetric $\epsilon$-tube respectively. It is because of the facts that $\alpha_i$ and $\beta_j$ also have to satisfy the KKT condition.

**Remark 4.** If proposed $\nu$-SVQR obtains the solution $(\bar{w}, \bar{b}, \bar{\epsilon})$ with parameter value $C'$, then $\epsilon$-SVQR model with parameters $\epsilon = \bar{\epsilon}$ and $C = C' \ast N$ will obtain the same solution $\bar{w}, \bar{b}$.

Obtaining the value of $\epsilon$ and $b$ :- At first we can obtain the value of the $\epsilon$ which is the effective width of the asymmetric tube. For this, we find out the data point which are lying on the upper and lower boundary of the $\epsilon$-tube using the preposition-5. For $0 < \alpha_i < \frac{C}{l} \tau$, we can obtain the upper width of the asymmetric tube using $y_i - (w^T \phi(x_i) + b)$. For $0 < \beta_j < \frac{C}{l} (1 - \tau)$, we can obtain the lower width of the asymmetric tube using $(w^T \phi(x_j) + b) - y_j$. But, the computation of final width of the asymmetric $\epsilon$-tube does not require the value of $b$ and can be obtained by

$$\epsilon = \left( \sum_{k=1}^{l} (\alpha_k - \beta_k) K(x_k, x_i) \right) + \left( \sum_{k=1}^{l} (\alpha_k - \beta_k) K(x_k, x_j) \right) - y_j)$$

where $0 < \alpha_i < \frac{C}{l} \tau$ and $0 < \beta_j < \frac{C}{l} (1 - \tau)$. (36)

After obtaining the value of $\epsilon$, we can obtain the value of $b$. For $0 < \alpha_i < \frac{C}{l} \tau$, we can obtain

$$b = y_i - \sum_{k=1}^{l} (\alpha_k - \beta_k) K(x_k, x_i) - (1 - \tau) \epsilon$$

(37)

For $0 < \beta_j < \frac{C}{l} (1 - \tau)$, we can also obtain

$$b = y_j - \sum_{k=1}^{l} (\alpha_k - \beta_k) K(x_k, x_j) + \tau \epsilon$$

(38)
In practice, we compute values of $b$ from equation (37) and (38) and use their average value as the final value of $b$. After computing the values of decision variables $\alpha$, $\beta$ and $b$ the quantile regression is estimated by

$$f_\tau(x) = \sum_{i=1}^{l} (\alpha_i - \beta_i) K(x, x_i) + b \quad (39)$$

Further, like $\epsilon$-SVR model described in (Gunn, [14]), if the kernel contains a bias term then, the $\nu$-SVQR dual problem (25) can be solved without equality constraint and the quantile regression function is simply estimated by

$$f_\tau(x) = \sum_{i=1}^{l} (\alpha_i - \beta_i) K(x, x_i) \quad (40)$$

4 Experimental Section

In this section, we shall empirically verify the claims made in this paper. For this, we first describe our experimental setup. We have performed all experiments with MATLAB 17.0 environment (http://in.mathworks.com/) on Intel i7 processor with 8.0 GB of RAM. The QPPs of proposed $\nu$-SVQR and $\epsilon$-SVQR has been solved by the quadprog function with interior-point convex algorithm available in the MATLAB 16.0 environment. For all of the experiments, we have used the RBF kernel function $\exp(-\frac{|x-y|^2}{q})$, where $q$ is the kernel parameter and quantile regression function is estimated by (39). The proposed $\nu$-SVQR model requires three parameters to be tuned namely RBF kernel parameter $q$, $C$ and $\nu$ whereas the $\epsilon$-SVQR model requires the tuning of parameters $q$, $C$ and $\epsilon$. All these parameters have been tuned using exhaustive search method (Hsu and Lin, [17]). The parameter $q$ and $C$ has been searched in the set $\{2^i : i = -15, -9, ......9, 15\}$.

4.1 Performance Criteria

For the evaluation of the efficacy of SVQR models, we have used some evaluation criteria which is also mentioned in (Xu Q et al., [18]). Given the training set $T = \{(x_i, y_i) : x_i \in \mathbb{R}^n, y_i \in \mathbb{R}, i = 1, 2, ..., l \}$ and true $\tau$-th conditional quantile function $Q_\tau(y/x)$, we list the evaluation criteria as follows.

(i) $RMSE$: It is Root Mean Square of Error.

It is given by $\sqrt{\frac{1}{l} \sum_{i=1}^{l} (Q_\tau(y_i/x_i) - f_\tau(x_i))^2}$.

(ii) $MAE$: It is Mean of the Absolute Error.

It is given by $\frac{1}{l} \sum_{i=1}^{l} |(Q_\tau(y_i/x_i) - f_\tau(x_i))|$

(iii) Error $E_\tau$: It is the measure which is used when the true quantile function is unknown. It is given by $E_\tau = |p_\tau - \tau|$, where $p_\tau = P(y_i \leq f_\tau(x_i))$ is the coverage probability. For the real world UCI datasets experiments, we
would be using this measure. We shall compute the coverage probability $p_\tau$ by obtaining the estimated $\tau$ value in 100 random trails.

(iv) \text{Sparsity}(u) = \frac{\#(u=0)}{\#(u)}$, where $\#(r)$ determines the number of the component of the vector $r$.

4.2 Artificial Datasets

We shall show different properties of proposed $\nu$-SVQR model and its advantages over $\epsilon$-SVQR model empirically. The best way to do this is to generate artificial datasets as actual true quantile can be easily computed for these datasets and unbiased comparisons can be made. We have generated the training set $T$ where $x_i$ is drawn from the univariate uniform distribution with $[-4,4]$. The response variable $y_i$ is obtained from polluting a nonlinear function of $x_i$ with different natures of noises in artificial datasets as follows.

AD1: $y_i = (1 - x_i + 2x_i^2)e^{-0.5x_i^2} + \xi$, where $\xi$ is from N(0, $\sigma$).

AD2: $y_i = (1 - x_i + 2x_i^2)e^{-0.5x_i^2} + \xi$, where $\xi$ is from U(a, b).

The true quantile function $Q_\tau(y_i/x_i)$ in these artificial datasets can be obtained as

$$y_i = (1 - x_i + 2x_i^2)e^{-0.5x_i^2} + F^{-1}_\tau(\xi),$$

where $F^{-1}_\tau(\xi)$ is the $\tau$th quantile of random error $\xi$. We have evaluated the SVQR models by generating 1000 testing points in each trails.

Experiment 1

Our first experiment will empirically verify the Preposition-6 of this paper. For this, we have generated 200 training data points of AD1 artificial datasets and obtain the numerical results for 10 random simulations. Table I shows the performance of the proposed $\nu$-SVQR model on several values of $\nu$ for $\tau=0.2, 0.5, 0.7$ and 0.8. Figure 2 shows the proposed $\nu$-SVQR model on several values of $\nu$ for $\tau=0.1, 0.3, 0.6$ and 0.9. Following observation can be easily drawn from numerical results listed in Table I and plots of Figure 2.

(a) Irrespective of $\tau$ values, as the value $\nu$ increases, the total width $\epsilon$ of asymmetric $\epsilon$-insensitive zone decreases.

(b) Irrespective of $\tau$ values, $\nu$ is upper bound on the fraction of of Errors.

(c) Irrespective of $\tau$ values, $\nu$ is lower bound on the fraction of of support vectors.

(d) RMSE and MAE obtained by proposed $\epsilon$-SVQR model also varies with the parameter $\nu$. 

Experiment 2

The second experiment has been performed with the varying number of training points of AD1 dataset for observing the asymptotic behavior of proposed $\nu$-SVQR model. For this experiment, we have fixed the $\nu = 0.8$ in proposed $\nu$-SVQR model. Table 2 results the numerical results obtained by proposed $\nu$-SVQR model on AD1 dataset with different size of training set. In this Table, ‘ratio’ is the ratio of training data points lying above and below the asymmetric $\epsilon$-insensitive tube. Following facts can be easily observed from the numerical results listed in Table 2.

(a) Irrespective of $\tau$ values, the fraction of support vectors and errors converges to the $\nu$ value in the proposed $\nu$-SVQR model. It has also been well illustrated by the plot in Figure 3(a) and 3(b). It is only because of fact that the probability of a training data point lying on boundaries of asymmetric $\epsilon$-insensitive tube vanishes, as the number of training point increases.

(b) Irrespective of $\tau$ values, the ratio of training data point lying above and below of the asymmetric $\epsilon$-tube converges to $\frac{(1-\tau)}{\tau}$. It has also been well illustrated by the plot in Figure 3(c) and 3(d). It means that the asymmetric $\epsilon$-insensitive zone used in proposed $\nu$-SVQR model is very suitable for handling quantile estimation problem.

(c) The resulting overall width of $\epsilon$-insensitive zone converges to a constant value in the proposed $\nu$-SVQR model.

(d) As the number of training points increases, there are more information available to the proposed $\nu$-SVQR model. It results in decrease in RMSE values obtained by proposed $\nu$-SVQR model.

Experiment 3

This experiment has been performed to show the capability of the proposed $\nu$-SVQR model to automate the control over accuracy. The proposed $\nu$-SVQR model has capability to automatically adjust the width of the asymmetric $\epsilon$-insensitive zone for efficient prediction. For fixed values of parameters with $\nu = 0.3$, we have simulated the proposed $\nu$-SVQR model on AD1 dataset with noise variance $\sigma = 0.2$ and $\sigma = 1$. Figure 4 shows the estimates obtained by proposed $\nu$-SVQR along with the $\epsilon$-insensitive zone for $\tau = 0.1$ and 0.9 at fixed value of $\nu = 0.3$. It can be observed that the proposed $\nu$-SVQR model can automatically adjust the width of the asymmetric $\epsilon$-insensitive zone according to the variance present in data for obtaining efficient estimates of quantiles.

Further, we have checked the performance of the proposed $\nu$-QSVR model on AD1 dataset with different noise variance $\sigma$. For this experiment, we have fixed the number of training data points to 500. The $\nu$ parameter in proposed $\nu$-QSVR was fixed to 0.4. Other parameters were also fixed. Table 3 lists numerical results obtained by the proposed $\nu$-SVQR model on AD1 dataset with different noise variance $\sigma$ for several $\tau$ values. Figure 5 illustrates the
numeral results listed in the Table 3 well for some $\tau$ values. Following things can be easily observed.

(a) Irrespective of values of $\tau$, as the noise variance $\sigma$ increases, the $\nu$-SVQR model accordingly increases the width of the asymmetric $\epsilon$-insensitive zone.
(b) Irrespective of values of $\sigma$, $\nu$ is the a upper bound on fraction of errors and lower bound on fraction of support vectors in proposed $\nu$-SVQR model.
(c) As the noise variance $\sigma$ increases, the RMSE obtained by $\nu$-SVQR model increases.

Experiment 4

As stated in Remark-4, the proposed $\nu$-SVQR model is similar to the $\epsilon$-SVQR model in the sense that any solution $(\hat{w}, \hat{b})$ obtained by the proposed $\nu$-SVQR can also be obtained by the $\epsilon$-SVQR. But, the proposed $\nu$-SVQR model has the capability of adjusting the $\epsilon$-insensitive zone according to the variance present in data. For realizing this direct benefit of proposed $\nu$-SVQR model over $\epsilon$-SVQR model, we perform the following experiment.

We generate 500 training data points of AD2 dataset where response points were polluted with noise from $U(-0.1, 0.1)$. For predicting the $\tau = 0.3$ quantile, we have tuned parameters of $\epsilon$-QSVR as well as proposed $\nu$-SVQR model. We have found that at $\epsilon = 0.1$ and $C = 2^9$, the $\epsilon$-QSVR model obtains the minimum RMSE 0.0057. The proposed $\nu$-SVQR model obtains the RMSE value 0.0056 with parameters $\nu = 0.5$ and $C = 2^9 \times 500$. The $\nu$-SVQR model obtains the asymmetric $\epsilon$ tube of width 0.0074. Now, with the same parameters setting in both $\epsilon$-QSVR and $\nu$-SVQR model, we increase the variance present in noise of AD2 dataset to $U(-5, 5)$. The $\epsilon$-SVQR model which has fixed $\epsilon$ value could obtain the RMSE 0.2168. But, the $\nu$-SVQR model automatically adjusts the width of asymmetric $\epsilon$-tube to 0.3734 and can obtain the RMSE value 0.1840.

Experiment 5

The above experiments are enough to empirically verify the claims made in this paper. But, we still want to check the performance of the $\nu$-SVQR model on real world data sets. For this, we have performed the experiments with the Servo (167×5) dataset which is taken from UCI repository (Blake, [4]). We have used 80% of this dataset for training the proposed $\nu$-SVQR model and rest of the data points were used for the testing. The 100 random trails have been used to obtain the Error $E_{\tau}$ and Sparsity. Table 4 shows the Error obtained by the proposed $\nu$-SVQR for different values of $\tau$ with different value of $\nu$. Table 5 shows the sparsity obtained by the proposed $\nu$-SVQR model for different value of $\nu$ with different $\tau$ values. It can be observed that irrespective of values of $\tau$, the sparsity decreases with increase in $\nu$ value.
5 Conclusions

We propose a novel $\nu$-Support Vector Quantile Regression ($\nu$-SVQR) model in this paper. There are several interesting properties of $\nu$-SVQR model which we have been briefly described and proved in this paper. Further, we have also empirically verified these proprieties by testing the proposed $\nu$-SVQR model on several artificial datasets as well as UCI dataset.
| $\nu$ | 0.050 | 0.100 | 0.150 | 0.200 | 0.250 | 0.300 | 0.350 | 0.400 | 0.450 | 0.500 | 0.550 | 0.600 | 0.650 | 0.700 | 0.750 | 0.800 | 0.850 | 0.900 | 0.950 | 1.000 |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| $\epsilon$ | 0.384 | 0.288 | 0.225 | 0.190 | 0.156 | 0.134 | 0.120 | 0.108 | 0.091 | 0.076 | 0.066 | 0.055 | 0.044 | 0.034 | 0.027 | 0.020 | 0.013 | 0.006 | 0.000 |
| SV | 0.000 | 0.125 | 0.210 | 0.230 | 0.290 | 0.335 | 0.380 | 0.420 | 0.480 | 0.520 | 0.580 | 0.630 | 0.650 | 0.710 | 0.725 | 0.770 | 0.800 | 0.850 | 0.925 | 1.000 |
| Error | 0.000 | 0.070 | 0.130 | 0.180 | 0.240 | 0.270 | 0.300 | 0.330 | 0.350 | 0.370 | 0.390 | 0.420 | 0.450 | 0.480 | 0.500 | 0.520 | 0.540 | 0.560 | 0.590 | 0.620 |
| RMSE | 0.160 | 0.072 | 0.030 | 0.035 | 0.044 | 0.062 | 0.065 | 0.065 | 0.068 | 0.069 | 0.066 | 0.064 | 0.063 | 0.064 | 0.065 | 0.066 | 0.067 | 0.068 | 0.068 | 0.069 |
| MAE | 0.127 | 0.059 | 0.037 | 0.032 | 0.034 | 0.041 | 0.043 | 0.043 | 0.043 | 0.043 | 0.043 | 0.043 | 0.043 | 0.044 | 0.044 | 0.044 | 0.044 | 0.044 | 0.044 | 0.044 |

| $\nu$ | 0.050 | 0.100 | 0.150 | 0.200 | 0.250 | 0.300 | 0.350 | 0.400 | 0.450 | 0.500 | 0.550 | 0.600 | 0.650 | 0.700 | 0.750 | 0.800 | 0.850 | 0.900 | 0.950 | 1.000 |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| $\epsilon$ | 0.000 | 0.120 | 0.110 | 0.230 | 0.210 | 0.250 | 0.300 | 0.330 | 0.350 | 0.370 | 0.390 | 0.420 | 0.450 | 0.480 | 0.500 | 0.520 | 0.540 | 0.560 | 0.580 | 0.600 |
| SV | 0.000 | 0.070 | 0.130 | 0.180 | 0.240 | 0.270 | 0.300 | 0.330 | 0.350 | 0.370 | 0.390 | 0.420 | 0.450 | 0.480 | 0.500 | 0.520 | 0.540 | 0.560 | 0.580 | 0.600 |
| Error | 0.000 | 0.070 | 0.130 | 0.180 | 0.240 | 0.270 | 0.300 | 0.330 | 0.350 | 0.370 | 0.390 | 0.420 | 0.450 | 0.480 | 0.500 | 0.520 | 0.540 | 0.560 | 0.580 | 0.600 |
| RMSE | 0.100 | 0.072 | 0.030 | 0.035 | 0.044 | 0.062 | 0.065 | 0.065 | 0.068 | 0.069 | 0.066 | 0.064 | 0.063 | 0.064 | 0.065 | 0.066 | 0.067 | 0.068 | 0.068 | 0.069 |
| MAE | 0.067 | 0.059 | 0.037 | 0.032 | 0.034 | 0.041 | 0.043 | 0.043 | 0.043 | 0.043 | 0.043 | 0.043 | 0.043 | 0.044 | 0.044 | 0.044 | 0.044 | 0.044 | 0.044 | 0.044 |

| $\nu$ | 0.050 | 0.100 | 0.150 | 0.200 | 0.250 | 0.300 | 0.350 | 0.400 | 0.450 | 0.500 | 0.550 | 0.600 | 0.650 | 0.700 | 0.750 | 0.800 | 0.850 | 0.900 | 0.950 | 1.000 |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| $\epsilon$ | 0.000 | 0.120 | 0.110 | 0.230 | 0.210 | 0.250 | 0.300 | 0.330 | 0.350 | 0.370 | 0.390 | 0.420 | 0.450 | 0.480 | 0.500 | 0.520 | 0.540 | 0.560 | 0.580 | 0.600 |
| SV | 0.000 | 0.070 | 0.130 | 0.180 | 0.240 | 0.270 | 0.300 | 0.330 | 0.350 | 0.370 | 0.390 | 0.420 | 0.450 | 0.480 | 0.500 | 0.520 | 0.540 | 0.560 | 0.580 | 0.600 |
| Error | 0.000 | 0.070 | 0.130 | 0.180 | 0.240 | 0.270 | 0.300 | 0.330 | 0.350 | 0.370 | 0.390 | 0.420 | 0.450 | 0.480 | 0.500 | 0.520 | 0.540 | 0.560 | 0.580 | 0.600 |
| RMSE | 0.100 | 0.072 | 0.030 | 0.035 | 0.044 | 0.062 | 0.065 | 0.065 | 0.068 | 0.069 | 0.066 | 0.064 | 0.063 | 0.064 | 0.065 | 0.066 | 0.067 | 0.068 | 0.068 | 0.069 |
| MAE | 0.067 | 0.059 | 0.037 | 0.032 | 0.034 | 0.041 | 0.043 | 0.043 | 0.043 | 0.043 | 0.043 | 0.043 | 0.043 | 0.044 | 0.044 | 0.044 | 0.044 | 0.044 | 0.044 | 0.044 |

Table 1: Performance of the proposed $\nu$-SVR model on AD1 dataset for different $\tau$ values.
| τ  | l    | 100  | 200  | 500  | 1000 | 3000 | 5000 |
|----|------|------|------|------|------|------|------|
|    | SV   | 0.85 | 0.85 | 0.81 | 0.81 | 0.80 | 0.80 |
|    | Error| 0.76 | 0.78 | 0.79 | 0.79 | 0.80 | 0.80 |
|    | Ratio| 4.13 | 4.13 | 9.72 | 9.72 | 9.20 | 9.17 |
|    | ϵ    | 0.04 | 0.05 | 0.05 | 0.04 | 0.05 | 0.05 |
|    | RMSE | 0.11 | 0.06 | 0.06 | 0.05 | 0.01 | 0.01 |

| 0.3 | l    | 100  | 200  | 500  | 1000 | 3000 | 5000 |
|-----|------|------|------|------|------|------|------|
|     | SV   | 0.85 | 0.82 | 0.82 | 0.81 | 0.80 | 0.80 |
|     | Error| 0.73 | 0.74 | 0.79 | 0.79 | 0.80 | 0.80 |
|     | Ratio| 2.65 | 2.56 | 2.41 | 2.38 | 2.36 | 2.35 |
|     | ϵ    | 0.04 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |
|     | RMSE | 0.19 | 0.06 | 0.06 | 0.06 | 0.01 | 0.01 |

| 0.7 | l    | 100  | 200  | 500  | 1000 | 3000 | 5000 |
|-----|------|------|------|------|------|------|------|
|     | SV   | 0.85 | 0.82 | 0.82 | 0.81 | 0.80 | 0.80 |
|     | Error| 0.73 | 0.76 | 0.79 | 0.79 | 0.80 | 0.80 |
|     | Ratio| 0.35 | 0.38 | 0.42 | 0.43 | 0.43 | 0.43 |
|     | ϵ    | 0.02 | 0.03 | 0.03 | 0.03 | 0.03 | 0.03 |
|     | RMSE | 0.07 | 0.06 | 0.06 | 0.06 | 0.01 | 0.01 |

| 0.9 | l    | 100  | 200  | 500  | 1000 | 3000 | 5000 |
|-----|------|------|------|------|------|------|------|
|     | SV   | 0.85 | 0.82 | 0.81 | 0.81 | 0.80 | 0.80 |
|     | Error| 0.76 | 0.77 | 0.79 | 0.79 | 0.80 | 0.80 |
|     | Ratio| 0.09 | 0.08 | 0.10 | 0.11 | 0.11 | 0.11 |
|     | ϵ    | 0.03 | 0.04 | 0.05 | 0.05 | 0.05 | 0.05 |
|     | RMSE | 0.13 | 0.06 | 0.07 | 0.03 | 0.02 | 0.02 |

Table 2: Performance of the proposed ν-SVQR with different size of training set.

Acknowledgment

We would like to acknowledge Ministry of Electronics and Information Technology, Government of India, as this work has been funded by them under Visvesvaraya PhD Scheme for Electronics and IT, Order No. Phd-MLA/4(42)/2015-16.

Conflict of Interest

We authors hereby declare that we do not have any conflict of interest with the content of this manuscript.
Fig. 3:
Asymptotic behavior of $\nu$-SVQR model for (a) $\tau = 0.1$ (b) $\tau = 0.3$ (c) $\tau = 0.1$ and (d) $\tau = 0.7$
| $\tau$ | $\sigma$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1 |
|-------|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|---|
|       | $\epsilon$ | 0.02 | 0.04 | 0.05 | 0.05 | 0.09 | 0.11 | 0.12 | 0.14 | 0.16 | 0.18 |
|       | Error    | 0.39 | 0.39 | 0.40 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 |
|       | SV      | 0.41 | 0.41 | 0.42 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.42 |
|       | RMSE    | 0.01 | 0.02 | 0.02 | 0.03 | 0.04 | 0.05 | 0.05 | 0.06 | 0.07 | 0.08 |
| 0.9   | $\sigma$ | 0.1  | 0.2  | 0.3  | 0.4  | 0.5  | 0.6  | 0.7  | 0.8  | 0.9  | 1  |
|       | $\epsilon$ | 0.01 | 0.03 | 0.04 | 0.06 | 0.07 | 0.09 | 0.10 | 0.12 | 0.13 | 0.15 |
|       | Error    | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 |
|       | SV      | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 |
|       | RMSE    | 0.00 | 0.01 | 0.02 | 0.02 | 0.03 | 0.03 | 0.04 | 0.04 | 0.05 | 0.05 |
| 0.7   | $\sigma$ | 0.1  | 0.2  | 0.3  | 0.4  | 0.5  | 0.6  | 0.7  | 0.8  | 0.9  | 1  |
|       | $\epsilon$ | 0.01 | 0.03 | 0.05 | 0.07 | 0.08 | 0.10 | 0.11 | 0.13 | 0.15 | 0.16 |
|       | Error    | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 |
|       | SV      | 0.42 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 |
|       | RMSE    | 0.01 | 0.01 | 0.02 | 0.02 | 0.03 | 0.03 | 0.03 | 0.04 | 0.04 | 0.04 |
| 0.5   | $\sigma$ | 0.1  | 0.2  | 0.3  | 0.4  | 0.5  | 0.6  | 0.7  | 0.8  | 0.9  | 1  |
|       | $\epsilon$ | 0.02 | 0.04 | 0.06 | 0.08 | 0.10 | 0.12 | 0.14 | 0.16 | 0.18 | 0.20 |
|       | Error    | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 | 0.39 |
|       | SV      | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 |
|       | RMSE    | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 |

Table 3: Performance of the proposed $\nu$-SVQR with different values of noise variance $\sigma$. 

| $\nu/\tau$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0.1       | 0.174 | 0.445 | 0.464 | 0.404 | 0.316 | 0.222 | 0.129 | 0.056 | 0.054 |
| 0.15      | 0.169 | 0.403 | 0.452 | 0.399 | 0.316 | 0.221 | 0.129 | 0.056 | 0.055 |
| 0.2       | 0.240 | 0.248 | 0.355 | 0.387 | 0.315 | 0.220 | 0.129 | 0.056 | 0.057 |
| 0.25      | 0.307 | 0.306 | 0.387 | 0.313 | 0.260 | 0.217 | 0.129 | 0.056 | 0.059 |
| 0.3       | 0.347 | 0.347 | 0.387 | 0.311 | 0.251 | 0.210 | 0.129 | 0.056 | 0.059 |
| 0.35      | 0.345 | 0.346 | 0.366 | 0.392 | 0.160 | 0.195 | 0.127 | 0.056 | 0.060 |
| 0.4       | 0.403 | 0.389 | 0.399 | 0.369 | 0.094 | 0.161 | 0.123 | 0.056 | 0.061 |
| 0.45      | 0.444 | 0.359 | 0.365 | 0.367 | 0.076 | 0.112 | 0.119 | 0.056 | 0.062 |
| 0.5       | 0.455 | 0.355 | 0.362 | 0.369 | 0.071 | 0.076 | 0.116 | 0.057 | 0.062 |
| 0.55      | 0.453 | 0.362 | 0.362 | 0.367 | 0.081 | 0.066 | 0.085 | 0.058 | 0.064 |
| 0.6       | 0.385 | 0.386 | 0.362 | 0.368 | 0.083 | 0.070 | 0.072 | 0.060 | 0.067 |
| 0.65      | 0.385 | 0.388 | 0.369 | 0.367 | 0.081 | 0.085 | 0.074 | 0.064 | 0.069 |
| 0.7       | 0.388 | 0.385 | 0.360 | 0.369 | 0.082 | 0.092 | 0.079 | 0.065 | 0.070 |
| 0.75      | 0.388 | 0.385 | 0.361 | 0.370 | 0.082 | 0.094 | 0.081 | 0.072 | 0.070 |
| 0.8       | 0.385 | 0.384 | 0.361 | 0.374 | 0.082 | 0.092 | 0.082 | 0.072 | 0.073 |
| 0.85      | 0.385 | 0.385 | 0.364 | 0.371 | 0.081 | 0.091 | 0.081 | 0.067 | 0.075 |
| 0.9       | 0.394 | 0.385 | 0.364 | 0.368 | 0.084 | 0.085 | 0.078 | 0.068 | 0.075 |
| 0.95      | 0.404 | 0.355 | 0.365 | 0.371 | 0.082 | 0.083 | 0.072 | 0.068 | 0.074 |
| 1.00      | 0.404 | 0.355 | 0.367 | 0.371 | 0.078 | 0.084 | 0.078 | 0.069 | 0.073 |

Table 4: Performance of the proposed $\nu$-SVQR model with different value of $\nu$ for different $\tau$ values.
Fig. 4: Automatic adjustment of the width of the $\epsilon$-insensitive zone in proposed $\nu$-SVQR model
Fig. 5: The proposed $\nu$-SVQR model with different noise variance $\sigma$ for (a) $\tau = 0.1$ (b) $\tau = 0.3$ (c) $\tau = 0.7$ and (d) $\tau = 0.9$

| $\nu/\tau$ | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 |
|-----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0.1       | 89.47 | 88.72 | 89.47 | 89.47 | 89.47 | 89.47 | 88.72 | 89.47 | 89.47 |
| 0.15      | 83.46 | 82.71 | 84.21 | 83.46 | 83.46 | 83.46 | 84.21 | 83.46 | 84.21 |
| 0.2       | 78.95 | 78.95 | 79.70 | 78.95 | 78.95 | 79.70 | 78.95 | 78.95 | 78.95 |
| 0.25      | 74.44 | 74.44 | 73.68 | 74.44 | 73.68 | 73.68 | 72.93 | 73.68 | 74.44 |
| 0.3       | 69.17 | 68.42 | 69.17 | 69.17 | 69.17 | 69.17 | 69.17 | 69.17 | 69.17 |
| 0.35      | 63.91 | 63.91 | 63.91 | 63.91 | 63.91 | 63.91 | 63.91 | 63.91 | 63.91 |
| 0.4       | 57.89 | 57.89 | 57.89 | 57.89 | 57.89 | 57.89 | 57.89 | 57.89 | 57.89 |
| 0.45      | 54.14 | 52.63 | 52.63 | 53.98 | 52.63 | 54.14 | 54.14 | 54.14 | 54.14 |
| 0.5       | 48.87 | 48.87 | 48.87 | 48.87 | 48.12 | 47.37 | 48.12 | 49.02 | 48.12 |
| 0.55      | 42.86 | 42.86 | 43.61 | 43.61 | 43.61 | 43.61 | 44.36 | 43.61 | 43.61 |
| 0.6       | 38.35 | 38.35 | 38.35 | 38.35 | 38.35 | 38.35 | 39.10 | 39.10 | 39.10 |
| 0.65      | 33.08 | 33.08 | 33.08 | 33.08 | 33.08 | 33.08 | 33.08 | 34.59 | 34.59 |
| 0.7       | 27.82 | 27.82 | 27.82 | 27.82 | 27.82 | 27.82 | 27.82 | 29.32 | 29.32 |
| 0.75      | 22.56 | 22.56 | 21.80 | 22.56 | 22.56 | 23.31 | 22.56 | 23.31 | 24.06 |
| 0.8       | 17.29 | 17.29 | 17.29 | 18.80 | 17.29 | 16.54 | 18.80 | 18.80 | 18.80 |
| 0.85      | 13.55 | 12.03 | 12.03 | 11.28 | 13.55 | 12.78 | 12.03 | 13.55 | 14.29 |
| 0.9       | 8.27  | 8.27  | 8.27  | 7.52  | 7.52  | 7.52  | 8.27  | 9.02  | 9.02  |
| 0.95      | 3.01  | 2.26  | 3.76  | 3.01  | 3.01  | 3.01  | 3.76  | 4.51  | 4.51  |
| 1         | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  | 0.00  |

Table 5: Sparsity obtained by the proposed $\nu$-SVQR model with different value of $\nu$ for different $\tau$ values.
Fig. 6: The $\epsilon$-SVQR model (left) fails to adjust the total width of the asymmetric $\epsilon$-insensitive tube with increase in noise variance. The proposed $\epsilon$-SVQR model (right) adjusts the total width of the asymmetric $\epsilon$-insensitive tube according to increase in noise variance and hence can obtain better estimate.
References

1. Koenker R, Bassett Jr G (1978) Regression quantiles. Econometrica, journal of the Econometric Society:33-50.
2. Koenker R (2005) Quantile Regression, Cambridge University Press.
3. Y Keming, RA Moyeed, Bayesian quantile regression, Statistics and Probability Letters 54.4(2001) pp 437-447.
4. Bosch RJ, Ye Y, Woodworth GG (1995) A convergent algorithm for quantile regression with smoothing splines. Computational Statistics and Data Analysis,19 pp 613-630.
5. Yu K, Jones MC (1998) Local linear quantile regression. Journal of the American statistical Association 93.441 pp 228-237.
6. Takeuchi I, Le QV, Scars T, Smola AJ (2006) Nonparametric quantile estimation. Journal of Machine Learning Research, 7, pp 1231-1264.
7. Vapnik V, Golovich S, Smola AJ (1997) Support vector method for function approximation, regression estimation and signal processing. Advances in neural information processing systems, 281-287.
8. Drucker H, Burges CJ, Kaufman L, Smola AJ, Vapnik V (1997) Support vector regression machines. Advances in neural information processing systems pp 155-161.
9. Vapnik V (1998) Statistical learning theory. Vol 1 New York Wiley.
10. Takeuchi I, Furuhashi T (2004) Non-crossing quantile regressions by SVM. IEEE International Joint Conference on Neural Networks Vol. 1, pp. 401-406.
11. Hu T, Xiang DH, Zhou DX (2012) Online learning for quantile regression and support vector regression. Journal of Statistical Planning and Inference, 142.12,3107-3122.
12. Seok KH, Cho D, Hwang C, Shim J (2010) Support vector quantile regression using asymmetric $\epsilon$-insensitive loss function. In Education Technology and Computer (ICETC), 2nd International Conference on Vol. 1 (2010), pp V1-438.
13. Park J, Kim J (2011). Quantile regression with an epsilon-insensitive loss in a reproducing kernel Hilbert space. Statistics and probability letters 81.1: 62-70.
14. Gunn S (1998) Support vector machines for classification and regression. ISIS technical report 14.1, pp 5-16.
15. Anand P, Rastogi R, Chandra S (2019). A new asymmetric $\epsilon$-insensitive pinball loss function based support vector quantile regression model, arXiv preprint arXiv:1908.06923.
16. SchÃüßler B, Smola AJ, Williamson RC, Bartlett PL (2000). New support vector algorithms. Neural computation, 12(5), 1207-1245.
17. Hsu CW, Lin CJ (2002). A comparison of methods for multi class support vector machines. IEEE Transaction on Neural Networks, 13 415-425.
18. Xu Q, Zhang J, Jiang C, Huang X, He Y (2015). Weighted quantile regression via support vector machine. Expert Systems with Applications, 42(13), pp 5441-5451.
19. Blake CL (1998) UCI repository of machine learning databases, irvine, university of california, , http://www.ics.uci.edu/~mlearn/MLRepository.html