Communications of the Korean Statistical Society
2010, Vol. 17, No. 2, 183–191

Support Vector Quantile Regression with Weighted Quadratic Loss Function

Jooyong Shim[a], Changha Hwang[1,b]

[a]Department of Applied Statistics, Catholic University of Daegu
[b]Department of Statistics, Dankook University

Abstract

Support vector quantile regression(SVQR) is capable of providing more complete description of the linear and nonlinear relationships among random variables. In this paper we propose an iterative reweighted least squares(IRWLS) procedure to solve the problem of SVQR with a weighted quadratic loss function. Furthermore, we introduce the generalized approximate cross validation function to select the hyperparameters which affect the performance of SVQR. Experimental results are then presented which illustrate the performance of the IRWLS procedure for SVQR.

Keywords: Support vector quantile regression, iterative reweighted least squares procedure, Kernel function, quadratic loss function, generalized approximate cross validation function.

1. Introduction

Quantile regression introduced by Koenker and Bassett (1978) is gradually involving into an ensemble of practical statistical methods for estimating and conducting inference about models for conditional quantile functions. Quantile regression is an increasingly popular method for estimating the quantiles of a distribution conditional on the values of covariates. Regression quantiles are robust against the influence of outliers and they give a more complete picture of the conditional distribution than a single estimate of the center. Just as classical linear regression methods based on minimizing sum of squared residuals enable one to estimate a wide variety of models for conditional mean functions, quantile regression methods offer a mechanism for estimating models for the conditional median function, and the full range of other conditional quantile functions. By supplementing the estimation of conditional mean functions with techniques for estimating an entire family of conditional quantile functions, quantile regression is capable of providing a more complete statistical analysis of the stochastic relationships among random variables. The introductions and current research areas of the quantile regression can be found in Koenker and Hallock (2001), Yu *et al.* (2003) and Kim *et al.* (2009).

The SVM, firstly developed by Vapnik (1995,1998), is being used as a new technique for regression and classification problems. SVM is gaining popularity due to many attractive features, and promising empirical performance. SVM was initially developed to solve classification problems but recently it has been extended to the domain of regression problems. SVM is based on the structural risk minimization(SRM) principle, which minimizes an upper bound on the expected risk unlike the

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (2009-0072101).

[1] Corresponding author: Professor, Department of Statistics, Dankook University, Gyeonggido 448-160, Korea.
E-mail: chwang@dankook.ac.kr

empirical risk minimization(ERM) minimizing the error on the training data. By minimizing this bound, high generalization performance can be achieved. In particular, for the SVM regression case SRM results in the regularized ERM with the e-insensitive loss function. The overviews and applications of SVM regression can be found in Vapnik (1995,1998), Smola and Schölkopf (1998) and Wang (2005).

The minimization problem associated with linear quantile regression is in essence the linear programming optimization problem, which is based on simplex algorithm or interior point algorithm. The current state of algorithms for nonlinear quantile regression is far less satisfactory. The widely used algorithm is interior point algorithm. Nonlinear quantile regression poses new algorithmic challenge. Refer to Koenker and Park (1996) and Koenker and Hallock (2001) for the algorithms. Training an SVM requires the solution to a quadratic programming(QP) optimization problem. But QP problem presents some inherent limitations which results in computational difficulty especially for the large data sets. Platt (1998), Flake and Lawrence (2002) developed the sequential minimal optimization algorithm which divides the QP problem into a series of small QP problems to avoid such computational difficulty. Perez-Cruz *et al.* (2000) proposed IRWLS algorithm for SVM regression(SVR) by transforming the Lagrangian function into sum of quadratic terms by defining associated weights of predicted errors.

In this paper we propose an IRWLS procedure to solve the QP problem of SVQR with a weighted quadratic loss function. The weighted quadratic loss function can provide the differentiability at 0, which enables to solve QP problem by IRWLS procedure. To select appropriate parameters for the achievement of high generalization performance, a commonly used method is minimizing the cross validation(CV) function. But selecting parameters using CV function is computationally formidable. Yuan (2006) proposed the generalized approximate cross validation(GACV) function for quantile spline estimation with modified check function. In SVQR using QP, GACV function or the generalized cross validation(GCV) function cannot be obtained by solving QP directly. We consider CV function which is different from that of Yuan (2006) and propose GCV function for the easy model selection.

The rest of this paper is organized as follows. In Section 2 we give a review of Support Vector Quantile regression(SVQR). In Section 3 we propose an IRWLS procedure for SVQR and present the model selection method using GCV function. In Section 4 we perform the numerical studies through examples. In Section 5 we give the conclusions.

2. Support Vector Quantile Regression

Quantile regression has long been studied in the literature. Most commonly used approach has been introduced by Koenker and Basset (1978). In this section we review the nonlinear quantile regression methods by implementing the idea of SVM.

Consider a random sample $\{\mathbf{x}_i, y_i\}_{i=1}^n$ with input vector $\mathbf{x}_i \in R^d$ including a constant 1 and output variable $y_i \in R$. Here the output variable y_i is related to the vector $\mathbf{x}_i$ of covariates. In the nonlinear quantile regression model the quantile function of the response y_i for a given $\mathbf{x}_i$ is assumed to be nonlinearly related to the input vector $\mathbf{x}_i \in R^d$. To allow for the nonlinear quantile regression, the input vectors $\mathbf{x}_i$ are nonlinearly transformed into a potentially higher dimensional feature space $\mathcal{F}$ by a nonlinear mapping function $\boldsymbol{\phi}(\cdot)$. The quantile function of the response y_i for a given $\mathbf{x}_i$ can be given as

$$q_\theta(\mathbf{x}_i) = \mathbf{w}_\theta^t \boldsymbol{\phi}(\mathbf{x}_i), \quad \text{for } \theta \in (0,1), \tag{2.1}$$

where $\mathbf{w}_\theta$ is the θ^{th} regression quantile. Here, similar to SVM for nonlinear regression, the nonlinear regression quantile estimator cannot be given in an explicit form since we use the kernel function of input vectors instead of the dot product of their feature mapping functions except for the identity feature mapping function such that $\boldsymbol{\phi}(\mathbf{x}) = \mathbf{x}$. Its estimator is defined as any solution to the optimization problem (Koenker and Basset, 1978),

$$\min \sum_{i=1}^{n} \rho_\theta(y_i - \mathbf{w}_\theta^t \boldsymbol{\phi}(\mathbf{x}_i)), \quad \text{for } \theta \in (0,1), \tag{2.2}$$

where $\rho_\theta(\cdot)$ is the check function defined as

$$\rho_\theta(r) = \theta r I(r \geq 0) + (\theta - 1) r I(r < 0).$$

Since quantile regression is in principle based on absolute deviation loss, to derive quantile regression using the idea of SVM, the procedures of the case $\epsilon = 0$ in a standard SVM is adopted. Then the quantile regression problem by the formulation for SVM can be expressed as,

$$\text{minimize } \frac{1}{2}\|\mathbf{w}_\theta\|^2 + C\sum_{i=1}^{n} \rho_\theta(y_i - \mathbf{w}_\theta^t \boldsymbol{\phi}(\mathbf{x}_i)), \quad \text{for } \theta \in (0,1). \tag{2.3}$$

The regularization parameter $C > 0$ determines the trade off between the flatness of quantile function estimate and the amount up to which deviations larger than 0 are tolerated.
By introducing slack variables ξ_i, ξ_i^*, we can rewrite (2.3) by following optimization problem,

$$\text{minimize } \frac{1}{2}\|\mathbf{w}_\theta\|^2 + C\sum_{i=1}^{n} (\theta\xi_i + (1-\theta)\xi_i^*, \quad \text{for } \theta \in (0,1), \tag{2.4}$$

$$\text{subject to } \begin{cases} y_i - \mathbf{w}_\theta^t \boldsymbol{\phi}(\mathbf{x}_i) \leq \xi_i, \\ \mathbf{w}_\theta^t \boldsymbol{\phi}(\mathbf{x}_i) - y_i \leq \xi_i^*, \\ \xi_i,\ \xi_i^* \geq 0. \end{cases}$$

The Lagrange function is constructed as follows:

$$\begin{aligned} L = {} & \frac{1}{2}\|\mathbf{w}_\theta\|^2 + C\sum_{i=1}^{n}(\xi_i + (1-\theta)\xi_i^*) - \sum_{i=1}^{n}\alpha_i(\xi_i - y_i + \mathbf{w}_\theta^t\boldsymbol{\phi}(\mathbf{x}_i)) \\ & - \sum_{i=1}^{n}\alpha_i^*(\xi_i^* + y_i - \mathbf{w}_\theta^t\boldsymbol{\phi}(\mathbf{x}_i)) - \sum_{i=1}^{n}(\eta_i\xi_i + \eta_i^*\xi_i^*). \end{aligned} \tag{2.5}$$

Notice that the positivity constraints $\alpha_i, \alpha_i^*, \eta_i, \eta_i^* \geq 0$ should be satisfied. After taking partial derivatives of Equation (2.5) with regard to the primal variables $(\mathbf{w}_\theta, \xi_i, \xi_i^*)$ and plugging them into Equation (2.5), the optimization problem with kernel function $K(\cdot,\cdot)$ is obtained as below,

$$\max_{\boldsymbol{\alpha},\boldsymbol{\alpha}^*} -\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(\mathbf{x}_i, \mathbf{x}_j) + \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)y_i \tag{2.6}$$

with constraints $\alpha_i \in [0, \theta C]$ and $\alpha_i^* \in [0, (1-\theta)C]$.

Solving the above optimization problem with the constraints determines the optimal Lagrange multipliers, α_i, α_i^*, the θ^{th} regression quantile estimators and the θ^{th} quantile function predictors of the input vector $\mathbf{x}$ are obtained, respectively as follows:

$$\mathbf{w}_\theta = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\boldsymbol{\phi}(\mathbf{x}_i) \text{ and } \hat{q}_\theta(\mathbf{x}) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(\mathbf{x}_i, \mathbf{x}). \tag{2.7}$$

Here, $\mathbf{w}_\theta$ and $\hat{q}_\theta(\mathbf{x})$ depend implicitly on θ through α_i and α_i^* depending on θ.

3. SVQR Using IRWLS Procedure

In this section we propose SVQR using IRWLS with weighted quadratic loss function. The check function $\rho_\theta(\cdot)$ used in SVQR can be seen as the weighted quadratic loss function such as

$$\rho_\theta(r) = wt(\theta)r^2, \tag{3.1}$$

where $wt(\theta) = \theta/|r|I(r \geq 0) + (1-\theta)/|r|I(r < 0)$ with an indicative function $I(\cdot)$. In IRWLS procedure with the weight obtained from previous step, we can use a differentiable weighted quadratic loss function instead of the check function which is not differentiable at 0. The representer theorem (Kimeldorf and Wahba, 1971) guarantees the minimizer of the optimization problem (2.3) to be $\hat{q}_\theta(\mathbf{x}) = K\boldsymbol{\beta}$ for some vector $\boldsymbol{\beta} = \boldsymbol{\alpha} - \boldsymbol{\alpha}^*$ and $C = \gamma/2$. Now the problem (2.3) becomes obtaining to minimize

$$L(\boldsymbol{\beta}) = \frac{1}{2}\boldsymbol{\beta}'K\boldsymbol{\beta} + \frac{\gamma}{2}\sum_{i=1}^{n} w_i(\theta)(y_i - K_i\boldsymbol{\beta})^2, \tag{3.2}$$

where $w_i(\theta) = \theta/|y_i - K_i\boldsymbol{\beta}|\ I(|y_i - K_i\boldsymbol{\beta}| \geq 0) + (1-\theta)/|y_i - K_i\boldsymbol{\beta}|\ I(|y_i - K_i\boldsymbol{\beta}| < 0)$ and K_i is the i^{th} row of K. Taking partial derivatives of (3.2) with regard to $\boldsymbol{\beta}$ leads to the optimal value of $\boldsymbol{\beta}$ to be the solution to

$$\mathbf{0} = K\boldsymbol{\beta} - \gamma K W_\theta \mathbf{y} + \gamma K W_\theta K\boldsymbol{\beta}. \tag{3.3}$$

Here W_θ is a diagonal matrix with the i^{th} diagonal element $w_i(\theta)$. The solution to (3.4) cannot be obtained in a single step since W_θ contains $\boldsymbol{\beta}$. Thus we need to apply IRWLS procedure which starts with initialized values of $\boldsymbol{\beta}$ as follows:

(a) Calculate W_θ with $\boldsymbol{\beta}$.

(b) Calculate $\boldsymbol{\beta}$ from $\boldsymbol{\beta} = (KW_\theta K + K/\gamma)^{-1}KW_\theta\mathbf{y}$.

(c) Iterate steps until convergence, where the proper criterion of convergence is found heuristically from the given data.

The functional structures of SVQR is characterized by hyperparameters - the regularization parameter and the kernel parameters. The cross validation(CV) technique used in SVR with the quadratic loss function cannot be used in SVQR since the check function used in SVQR is not differentiable as the quadratic loss function. To select the parameters of SVQR using IRWLS we consider the cross validation(CV) function as follows:

$$\mathrm{CV}(\boldsymbol{\lambda}) = \frac{1}{n}\sum_{i=1}^{n} w_i(\theta)\left(y_i - \hat{q}_\theta^{(-i)}(\mathbf{x}_i)\right)^2, \tag{3.4}$$

instead of the cross validation(CV) function used in SVQR using QP with check function, which is,

$$\mathrm{CV}(\lambda) = \frac{1}{n}\sum_{i=1}^{n}\rho_\theta\left(y_i - \hat{q}_\theta^{(-i)}(\mathbf{x}_i)\right), \tag{3.5}$$

where λ is the set of parameters and $\hat{q}_\theta^{(-i)}(\mathbf{x}_i)$ is the quantile function estimated without i^{th} observation. Since for each candidates of parameters, $\hat{q}_\theta^{(-i)}(\mathbf{x}_i)$ for $i = 1,\ldots,n$, should be evaluated, selecting parameters using CV function is computationally formidable. By leaving-out-one lemma (Craven and Wahba, 1979) ,

$$\left(y_i - \hat{q}_\theta^{(-i)}(\mathbf{x}_i|\lambda)\right) - (y_i - \hat{q}_\theta(\mathbf{x}_i|\lambda)) = \hat{q}_\theta(\mathbf{x}_i|\lambda) - \hat{q}_\theta^{(-i)}(\mathbf{x}_i|\lambda) \approx \frac{\partial\hat{q}_\theta(\mathbf{x}_i|\lambda)}{\partial y_i}\left(y_i - \hat{q}_\theta^{(-i)}(\mathbf{x}_i|\lambda)\right),$$

we have

$$\left(y_i - \hat{q}_\theta^{(-i)}(\mathbf{x}_i|\lambda)\right) \approx \frac{y_i - \hat{q}_\theta(\mathbf{x}_i|\lambda))}{1 - \partial\hat{q}_\theta(\mathbf{x}_i|\lambda)/\partial y_i}.$$

Then the ordinary cross validation(OCV) function can be obtained as

$$\mathrm{OCV}(\lambda) = \frac{1}{n}\sum_{i=1}^{n} w_i(\theta)\left(\frac{y_i - \hat{q}_\theta(\mathbf{x}_i|\lambda)}{1 - \partial\hat{q}_\theta(\mathbf{x}_i|\lambda)/\partial y_i}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} w_i(\theta)\left(\frac{y_i - \hat{q}_\theta(\mathbf{x}_i|\lambda)}{1 - h_{ii}}\right)^2, \tag{3.6}$$

where $H = K(KW_\theta K + K/\gamma)^{-1}KW_\theta$ is a hat matrix such that $\hat{q}_\theta(\mathbf{x}_i|\lambda)) = H\mathbf{y}$ and h_{ii} is the i^{th} diagonal element of H. Replacing h_{ii} by their average $\mathrm{tr}(H)/n$, the generalized cross validation(GCV) function can be obtained as

$$\mathrm{GCV}(\lambda) = \frac{n}{n - \mathrm{tr}(H)}\sum_{i=1}^{n} w_i(\theta)\,(y_i - \hat{q}_\theta(\mathbf{x}_i|\lambda))^2. \tag{3.7}$$

4. Numerical Studies

In this section, we illustrate the performance of the proposed quantile regression estimation using IRWLS through the simulated data sets and motorcycle data set (Härdle, 1989). We compare the performance of the proposed quantile regression estimation(SVQR using IRWLS) with that of Takeuci *et al.* (2006) which uses the quadratic programming(SVQR using QP). The numerical studies are conducted in MATLAB environment.

Example 1. We generate one training data set of size 300 and 100 test data sets of size 300 in a similar manner to Cawley *et al.* (2004). The univariate input observations x's are equally spaced ranging from 0 to π, the corresponding responses y's are drawn from a univariate normal distribution with mean and variance that vary smoothly with x as follows:

$$y = \sin\left(\frac{5x}{2}\right)\sin\left(\frac{3x}{2}\right) + \sigma(x)\epsilon, \quad \text{with } \sigma(x) = \sqrt{\frac{1}{100} + \frac{1}{4}\left\{1 - \sin\left(\frac{5x}{2}\right)\right\}^2}, \ \epsilon \sim N(0,1).$$

The Gaussian kernel function is utilized in this example, which is

$$K(\mathbf{x}_1, \mathbf{x}_2) = \exp\left(-\frac{1}{\sigma^2}\|\mathbf{x}_1 - \mathbf{x}_2\|^2\right),$$

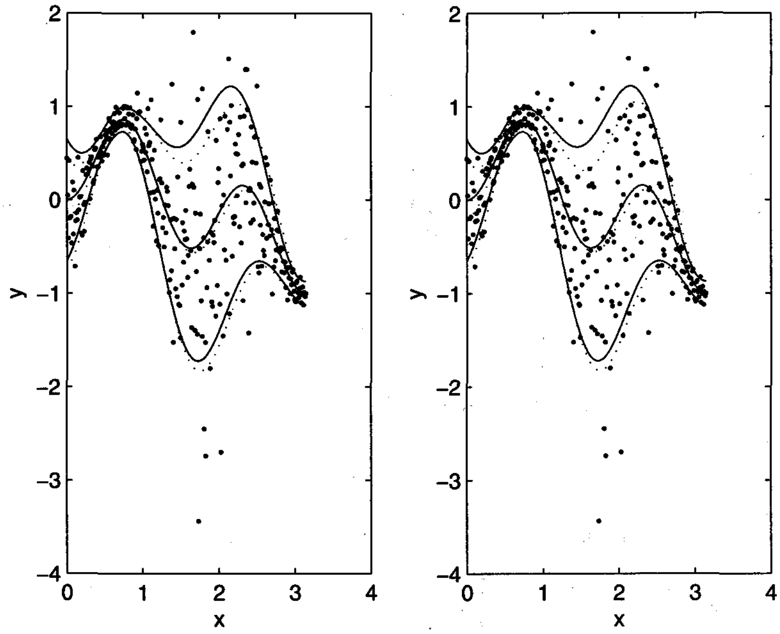

Figure 1: *An illustration of the proposed SVQR using IRWLS(Left) and QP(Right) for the training data set of size 300 of Example 1. True quantile regression functions(solid lines) and the estimated quantile regression functions(dotted lines) are superimposed on the scatter plot.*

where $\mathbf{x} = (1, x_i)$ for $i = 1, 2$. Figure 1 shows a family of quantile functions estimated by SVQR using IRWLS and QP for the training data set. The estimated quantile regression functions for $\theta = 0.1, 0.5, 0.9$ are superimposed on the scatter plot. In SVQR using IRWLS the values of γ and σ^2 are chosen by GCV function (3.7) such as (300, 1.5) for $\theta = 0.1$, (400, 2) for $\theta = 0.5$ and (400, 2) for $\theta = 0.9$. In SVQR using QP the values of and are chosen by CV function (3.5) such as (400, 1.5) for $\theta = 0.1$, (300, 2) for $\theta = 0.5$ and (300, 1.5) for $\theta = 0.9$. As seen from Figure 1, in both procedures the three estimated quantile regression functions reflect well the heteroscedastic structure of the error terms. They have their (local) minima and (local) maxima at different values. For example, the 0.1^{th}, 0.5^{th} and 0.9^{th} quantile regression functions have maxima at $x = 0.75$, 0.75 and 2.15, respectively, and minima at 1.75, π and π, respectively. To illustrate the prediction performance of SVQR using IRWLS, we compare it with SVQR using QP via 100 data sets, where the mean squared error(MSE) is used as the estimation performance measure defined by

$$\text{MSE} = \frac{1}{100}\sum_{i=1}^{100}(q_\theta(x_i) - \hat{q}_\theta(x_i))^2 .$$

The averages of 100 MSEs from SVQR using IRWLS and QP are obtained as 0.0173 and 0.0174 for $\theta = 0.1$, 0.0082 and 0.0084 for θ =0.5, 0.0154 and 0.0160 for θ =0.9, respectively. We can see that both procedures have almost same estimation performance for Example 1.

Example 2. We generate one training data set of size 300 and 100 test data sets of size 300 in a similar manner to Cawley *et al.* (2004). The univariate input observations x's are equally spaced

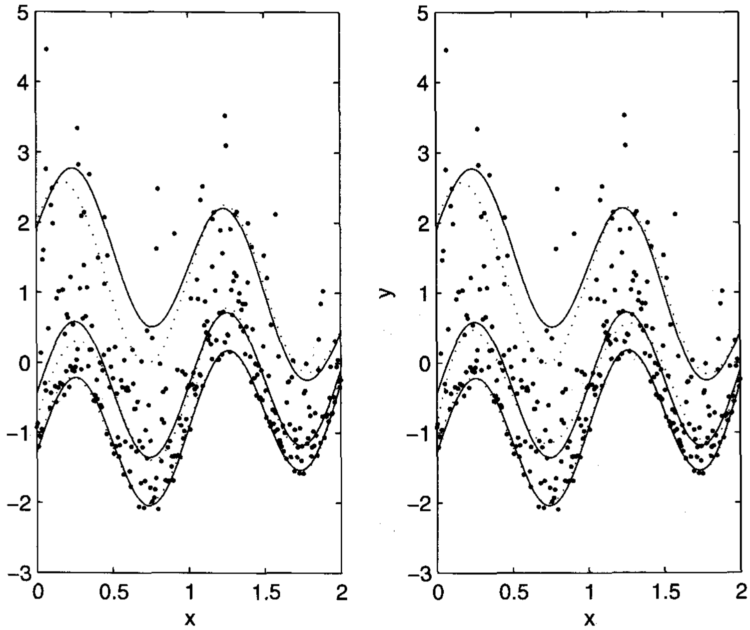

Figure 2: *An illustration of the proposed SVQR using IRWLS(Left) and QP(Right) for the training data set of size 300 of Example 2. True quantile regression functions(solid lines) and the estimated quantile regression functions(dotted lines) are superimposed on the scatter plot.*

ranging from 0 to 2, the corresponding responses y's are drawn from a univariate χ^2 distribution with mean and variance that vary smoothly with x as follows:

$$y = \sin(2\pi x) + \sigma(x)\epsilon, \quad \text{with } \sigma(x) = \sqrt{\frac{2.1 - x}{4}}, \ \epsilon \sim \chi^2_{(2)} - 2.$$

Here $\chi^2_{(2)}$ is the chi-squared distribution with degree of freedom 2. The Gaussian kernel function is utilized in this example as in Example 1. Figure 2 shows a family of quantile functions estimated by SVQR using IRWLS and QP for the training data set. The estimated quantile regression functions for $\theta = 0.1, 0.5, 0.9$ are superimposed on the scatter plot. In SVQR using IRWLS the values of and are chosen by GCV function (3.7) such as (100, 0.5) for $\theta = 0.1$, (400, 0.5) for $\theta = 0.5$ and (400, 0.5) for $\theta = 0.9$. In SVQR using QP the values of γ and σ^2 are chosen by CV function (3.5) such as (100, 0.5) for $\theta = 0.1$, (200, 1) for $\theta = 0.5$ and (100, 0.5) for $\theta = 0.9$. As seen from Figure 2, in both procedures the three estimated quantile regression functions reflect well the heteroscedastic and nonnormal structure of the error terms. To illustrate the prediction performance of SVQR using IRWLS, we compare it with SVQR using QP via 100 data sets, where the mean squared error(MSE) is used as the estimation performance measure. The averages of 100 MSEs from SVQR using IRWLS and QP are obtained as 0.0034 and 0.0033 for $\theta = 0.1$, 0.0303 and 0.0446 for $\theta = 0.5$, 0.2452 and 0.2162 for $\theta = 0.9$, respectively. We can see that both procedures have almost same estimation performance for Example 2.

Example 3. In this example we consider the motorcycle data, which have been widely used to demonstrate the performance of nonparametric quantile regression methods. The data were collected

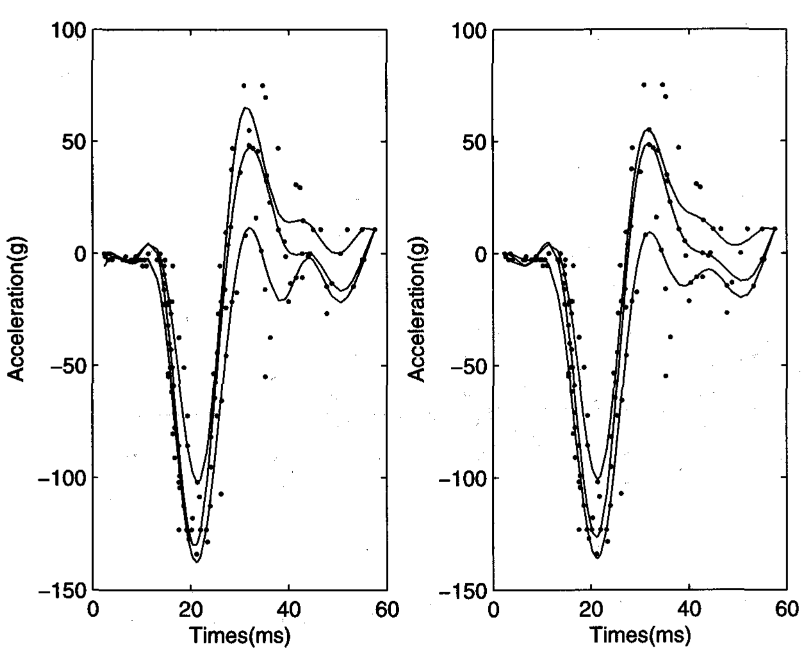

Figure 3: *An illustration of the proposed SVQR using IRWLS(Left) and QP(Right) for the mortorcycle data set of Example 3. The estimated quantile regression functions are superimposed on the scatter plot.*

performing crash tests with dummies sitting on motorcycles. The head acceleration(y) of the dummies (in g) was recorded a certain time (measured in milliseconds(x)) after they had hit a wall. The estimated quantile regression functions for $\theta = 0.25, 0.5, 0.75$ are superimposed on the scatter plot in Figure 1. In SVQR using IRWLS the values of and are chosen by GCV function (3.7) such as (100, 0.5) for $\theta = 0.25$, (200, 0.5) for $\theta = 0.5$ and (200, 0.75) for $\theta = 0.75$. In SVQR using QP the values of γ and σ^2 are chosen by CV function (3.5) such as (100, 0.75) for $\theta = 0.25$, (200, 0.75) for $\theta = 0.5$ and (10, 0.5) for $\theta = 0.75$ As seen from Figure 3, as x increases the variance of y increases when $x < 33$ and decreases when $x > 33$. We can see that both procedures have almost same estimation performance for the motorcycle data.

5. Conclusions

In this paper, we dealt with estimating the nonlinear quantile regression function by SVQR using IRWLS procedure with a weighted quadratic loss function and obtained GCV function for the proposed method, which provides faster computation time than SVQR using QP. Through the examples we showed that the proposed method derives the satisfying solutions. We also found that SVQR using IRWLS procedure is much faster than SVQR using QP, which implies that the proposed method is appropriate for the large training data sets.

References

Cawley, G. C., Talbot, N. L. C., Foxall, R. J., Dorling, S. R. and Mandic, D. P. (2004). Heteroscedastic Kernel ridge regression, *Neurocomputing*, **57**,105-124.

Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation, *Numerical Mathematics*, **31**, 377-403

Flake, G. W. and Lawrence, S. (2002). Efficient SVM regression training with SMO, *Machine Learning*, **46**, 271–290.

Härdle, W. (1989). *Applied Nonparametric Regression*, Cambridge University Press, Cambridge.

Kim, Y., Shim, J., Lee, J. T. and Hwang, C. (2009). Combination of value-at-risk models with support vector machine, *Communications of the Korean Statistical Society*, **16**, 791–801.

Kimeldorf, G. S. and Wahba, G. (1971). Some results on Tchebycheffian spline functions, *Journal of Mathematical Analysis and its Applications*, **3**, 82–95.

Koenker, R. and Bassett, G. (1978). Regression quantile, *Econometrica*, **46**, 33–50

Koenker, R. and Hallock, K. F. (2001). Quantile regression, *Journal of Economic Perspectives*, **40**, 122–142.

Koenker, R. and Park, B. J. (1996). An interior point algorithm for nonlinear quantile regression, *Journal of Econometrics*, **71**, 265-283.

Perez-Cruz, F., Navia-Vazquez, A., Alarcon-Diana, P. L. and Artes- Rodriguez, A. (2000). An IRWLS procedure for SVR, In *Proceedings of European Association for Signal Processing, EUSIPO 2000, Tampere*, Finland.

Platt, J. (1998). Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines, Microsoft Research Technical Report MSR-TR-98-14.

Smola, A. J. and Schölkopf, B. (1998). On a Kernel-based method for pattern recognition, regression, *Approximation and Operator Inversion Algorithmica*, **22**, 211–231.

Takeuchi, I., Le, Q. V., Sears, T. D. and Smola, A. J. (2006). Nonparametric quantile estimation, *Journal of Machine Learning Research*, **7**, 1231–1264.

Vapnik, V. N. (1995). *The Nature of Statistical Learning Theory*, Springer, New York.

Vapnik, V. N. (1998). *Statistical Learning Theory*, John Wiley, New York.

Wang, L. (2005). *Support Vector Machines: Theory and Application*, Springer, Berlin Heidelberg, New York.

Yu, K., Lu, Z. and Stander, J. (2003). Quantile regression: Applications and current research area, *The Statistician*, **52**, 331–350.

Yuan, M. (2006). GACV for quantile smoothing splines, *Computational Statistics and Data Analysis*, **50**, 813–829.

Received October 2009; Accepted November 2009