Power Weighted Quantile Regression and Its Application

Xue J. Ma[1*] and Feng X. He[2]

[1]Renmin University of China
[2]North China Electric Power University

Abstract: In the paper, we propose power weighted quantile regression(PWQR), which can reduce the effect of heterogeneous of the conditional densities of the response effectively and improve efficiency of quantile regression). In addition to PWQR, this article also proves that all the weighting of those that the actual value is less than the estimated value of PWQR and the proportion of all the weighting is very close to the corresponding quantile. At last, this article establishes the relationship between Geomagentic Indices and GIC. According to the problems of power system security operation, we make GIC risk value table. This table can have stronger practical operation ability, can provide power system security operation with important inferences.

Key words: Quantile regression, Power Weighted Quantile Regression, GIC, Geomagentic Indices

1. Introduction

Weighted quantile regression is proposed by Koenker(2005). When the conditional densities of the response are heterogeneous, he maintained weighted quantile regression might lead to efficiency improvements. It is considered, but reweighting based on estimated densities is somewhat difficult, that is the estimation of the weights is hard to do well, so it isn't usually done. Often researchers prefer to try to correct sparsity function or bootstrap (Hendricks and Koenker 1991,Parzen et al 1994) for non iid error behavior. If the covariance structure were known, weighted quantile regression would be more efficient. With regard to research weighted quantile regression, Taylor(2008) proposed exponentially weighted quantile regression and applied in Value at Risk and expected shortfall. But if data exists node, exponentially weighted effect would be meaningless. For example, when one independent variable value corresponding to more than one dependent variable values, the method of exponentially weight is clearly unreasonable. Because the time sequence of the observed values are different, weights are different. So that this paper proposes a new weighting method, that is Power Weighted Quantile Regression.

---

[*] Corresponding author.

2. Power Weighted Quantile Regression

In the light of the development of Quantile regression, we propose power weighted quantile regression:

$$\min_{\beta \in R^p} \sum_{i=1}^{n} ||x_i||^{\alpha} \rho_{\tau} \left(y_i - x_i^T \beta\right) \tag{1}$$

where $\| x \|$ is norm. Quantile regression is the special case of power weighted quantile regression. Because when $\alpha = 0$, power weighted quantile regression is quantile regression.

Theorem 1 If X contains in quantile regression model $Q_{\tau}(y/x) = x^T \beta$, then for any $\hat{\beta}$, solving (1), we have

$$\frac{\sum_{i=1}^{n}||x_i||^{\alpha}I(y_i < x_i^T\hat{\beta})}{\sum_{i=1}^{n}||x_i||^{\alpha}} \leq \tau \leq \frac{\sum_{i=1}^{n}||x_i||^{\alpha}I(y_i < x_i^T\hat{\beta}) + \sum_{i=1}^{n}||x_i||^{\alpha}I(y_i = x_i^T\hat{\beta})}{\sum_{i=1}^{n}||x_i||^{\alpha}} \tag{2}$$

$$\frac{\sum_{i=1}^{n}||x_i||^{\alpha}I(y_i > x_i^T\hat{\beta})}{\sum_{i=1}^{n}||x_i||^{\alpha}} \leq 1 - \tau \leq \frac{\sum_{i=1}^{n}||x_i||^{\alpha}I(y_i > x_i^T\hat{\beta}) + \sum_{i=1}^{n}||x_i||^{\alpha}I(y_i = x_i^T\hat{\beta})}{\sum_{i=1}^{n}||x_i||^{\alpha}} \tag{3}$$

Proof. See Appendix.

The theorem indicates that the estimator of PWQR essentially partitions divided the y_i observations into three parts $y_i < x_i^T\hat{\beta}$, $y_i = x_i^T\hat{\beta}$, $y_i < x_i^T\hat{\beta}$. The sum of the weights on $y_i < x_i^T\hat{\beta}$, as a proportion of the sum of all the weights, is close to τ. The proportion is not exactly equal to τ, because $y_i = x_i^T\hat{\beta}$ may exist. That is to say, given τ, the weighted proportion of those observations, which underlie quantile curve, is close to τ. Similarly, the sum of the weights on $y_i > x_i^T\hat{\beta}$, as a proportion of the sum of all the weights, is close to $1 - \tau$.

In (1), we must deal with the important α selection problem, as the quality of the linear estimates depends sensitively on the choice of it. A convenient and effective method is that we can use goodness of fit, proposed by Koenker and Machado(1999). Consider the quantile regression model

$$Q_T(y_i, x_i) = x_i^T \beta$$

which we can partition as

$$Q_T(y_i, x_i) = x_{i1}^T \beta_1 + x_{i2}^T \beta_2$$

and let $\hat{V}(\tau, \alpha) = \min_{\beta \in R^p} \sum_{i=1}^{n} ||x_i||^{\alpha} \rho_{\tau}(y_i - x_i^T \beta)$, $\tilde{V}|(\tau, \alpha) = \min_{\beta \in R^{p-q}} \sum_{i=1}^{n} ||x_i||^{\alpha} \rho_{\tau}(y_i - x_{1i}^T \beta_1)$, so the goodness of fit is

$$R^1(\tau, \alpha) = 1 - \frac{\hat{V}(\tau, \alpha)}{\tilde{V}(\tau, \alpha)}$$

Conventionally, x_{i1} including only an "intercept" parameter (x_{i1}=1) yields the R^1 usually reported. $R^1(\tau, \alpha)$ measures the relative success of the corresponding quantile regression

models at a specific quantile with reference to an appropriately weighted sum of absolute residuals, which can assess goodness of fit for quantile regression models. $R^1(\tau, \alpha)$ lies between 0 and 1.Generally speaking, it's better to have a larger $R^1(\tau, \alpha)$. The optimal α can be selected by maximizing $R^1(\tau, \alpha)$, that is $\alpha_{\text{opt}}(\tau) = arg \max_a R^1(\tau, \alpha)$

3. Data

Geomagentic Indices depicts the overall intensity of geomagenitic disturbance in a time slot or the index of certain geomagenitic disturbance, such as frequently-used indexes of K, Kp, ap and so on, of which the index K is employed to describe geomagenitic disturbance intensity of a single geomagnetic observatory in a time slot of three hours, which means that index K could only describe the geomagentic activity in certain district, not the globe. In order to present the geomagentic activity intensity of the globe, it needs a new index Kp by standardizing and then averaging 12 geomagnetic observatories from global Geomagnetic station network. Because of the nonlinear relationship between Kp and geomagenitic disturbance, it is transformed into linear relationship to get the index Kp. The size of a geomagnetic storm is classified as moderate, intense and super-storm according to the size of Geomagentic Indices. Generally speaking, the higher of the geomagentic storm intensity, the larger of GIC caused by geomagentic storm, which demonstrates the close relation between GIC and Geomagentic Indices.

As for the evaluation of GIC, there are basically two models: physical model and statistical model, of which the later one is the focus of this article. Trichtchenko and Boteler(2004) from Canada with others has established class linear regression of GIC using ap. He and Ma(2011)used quantile regression. This article tries to improve the quantile regression and make advantage of power weighted quantile regression. Data sources: the data of GIC is based on the monitoring of Guangzhou lingao nuclear power plant when geomagnetic storm took place in 2004 to 2005, and the data of ap come from Center for Space Environment Research and Forecast.

3.1 Comparing quantile regression and power weighted quantile regression

As for GIC and ap, we need to establish unitary quantile regression model and unitary power weighted quantile regression model respectively. What's more, we also need to give the regression coefficient when the quantile is0.05,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,0.95 respectively. According to goodness of fit, the corresponding selection of α is 0.0,-0.4,-0.1,-2.6,-0.4,-0.4,-0.9,-0.4,-0.9,-1.4, 0.0.

From the table 1, the goodness of fit of PWQR is larger than QR except $\alpha = 0$. The all the coefficient given by PWQR is remarkable at the level of 0.01, with only part of the coefficient given by QR is remarkable at the level of 0.01. When tau equals to 0.3, 0.4, 0.5, 0.6 respectively, the significance of QR fails the test, which means that it is not notable, while PWQR is all very notable. Except for the value of PWQR and QR are the same, the fact that the

coefficient before ap of PWQR are all smaller than those of QR and constant term of PWQR are all bigger than those of QR. From the aspect of the significance of regression coefficient and goodness of fit, it is easy to judge that the PWQR model is superior to the quantile regression model.

3.2 Assessing GIC by PWQR

From the above section, we know the coefficient when tau=0.05, 0.1, 0.3, 0.5, 0.7, 0.9, 0.95 respectively. As follows:

$$Q_{0.05}(GIC|_{ap}) = 0.546 \ + \ 0.015ap$$
$$Q_{0.1}(GIC|_{ap}) = 0.573 \ + \ 0.018ap$$
$$Q_{0.3}(GIC|_{ap}) = 0.946 \ + \ 0.029ap$$
$$Q_{0.5}(GIC|_{ap}) = 1.006 \ + \ 0.048ap$$
$$Q_{0.7}(GIC|_{ap}) = 1.459 \ + \ 0.080ap$$
$$Q_{0.9}(GIC|_{ap}) = 5.703 \ + \ 0.144ap$$
$$Q_{0.95}(GIC|_{ap}) = 6.942 \ + \ 0.210ap$$

Table 1: Rregression coefficient of PWQR and QR.

tau	Coef.	PWQR				QR			
		Est.	Sd	p.value	R^1	Est.	Sd	p.value	R^1
0.05	Intercept	0.546	0.190	0.005	0.095	0.546	0.190	0.005	0.095
	ap	0.015	0.002	0.000		0.015	0.002	0.000	
0.1	Intercept	0.573	0.146	0.000	0.114	0.525	0.213	0.016	0.112
	ap	0.018	0.003	0.000		0.018	0.002	0.000	
0.2	Intercept	0.467	0.202	0.023	0.141	0.356	0.179	0.049	0.141
	ap	0.028	0.002	0.000		0.029	0.002	0.000	
0.3	Intercept	0.946	0.001	0.000	0.199	0.645	0.339	0.060	0.159
	ap	0.029	0.004	0.000		0.034	0.003	0.000	
0.4	Intercept	0.900	0.272	0.001	0.202	0.518	0.309	0.097	0.201
	ap	0.040	0.005	0.000		0.049	0.003	0.000	
0.5	Intercept	1.006	0.287	0.001	0.243	0.723	0.387	0.065	0.231
	ap	0.048	0.005	0.000		0.053	0.003	0.000	
0.6	Intercept	1.617	0.177	0.000	0.283	0.752	0.463	0.108	0.265
	ap	0.048	0.011	0.000		0.068	0.004	0.000	
0.7	Intercept	1.459	0.628	0.022	0.315	1.208	0.530	0.025	0.301
	ap	0.080	0.012	0.000		0.082	0.005	0.000	
0.8	Intercept	4.259	0.497	0.000	0.379	1.649	0.594	0.007	0.326
	ap	0.072	0.030	0.017		0.096	0.005	0.000	
0.9	Intercept	5.703	0.150	0.000	0.515	5.733	2.676	0.035	0.371
	ap	0.144	0.021	0.000		0.142	0.024	0.000	
0.95	Intercept	6.942	1.398	0.000	0.344	6.942	1.398	0.000	0.344
	ap	0.210	0.012	0.000		0.210	0.012	0.000	

By the expressions, we could see that regression constant given by power weighted quantile and the coefficient of ap are inclined to increase with the increase of tau, of which tau=0.05 represents the low tail distribution of GIC, tau=0.3 represents mid low tail distribution, tau=0.5 represents middle information of GIC, tau=0.7 represents mid high tail information and tau=0.9 or 0.95 represents the high tail information. So by the analysis, we could find that PWQR is able to depict the overall information of GIC.

In Figure 1, from the bottom to the top, these straight lines represents regression lines when quantiles are 0.05, 0.1, 0.3, 0.5, 0.7, 0.9, 0.95 respectively. Furthermore, with the growing of ap, GIC tends to go up, and when ap < 110, GIC is densely distributed in the low position, with a few in higher position. At this point, we could use different quantile to fit the value of GIC, of which for certain ap, we could get different GIC fitted value of different tau, rendering the value given by quantile regression being able to reflect the overall information of GIC. For example,
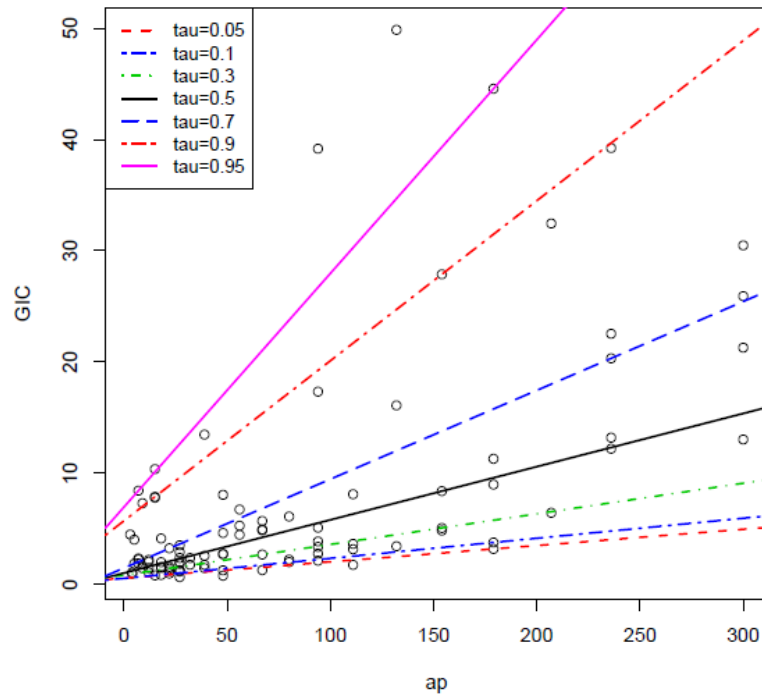
Figure 1: Fitted lines of PWQR

when ap=179, the actual value are 3.172, 3.755, 8.966, 11.282 or 44.566. The fitted value of different tau in the models is as following: For the actual value, we could find the fitted value of tau by quantile regression, such as GIC=3.75, tau=0.1 is used to fit. In practice, our concern is those GIC that could pose damage to electric grid, in other words, the high tail information of GIC distribution; or the middle level of GIC, and that is to say the PWQR when tau=0.5

Table 2: Fitted values of different quantile

tau	0.05	0.1	0.3	0.5	0.7	0.9	0.95
PWQR	3.172	3.749	5.754	9.57	15.762	31.476	44.566

According to PWQR, we get the fitted value of high tail and middle GIC when *ap* > 111 (table 3). What we should highlight is that we could not only get the fitted value of GIC and also its probability by PWQR. For example, when *ap*=179, we are convinced that the probability of GIC that is no less than 9.57 is 50% and of GIC that is no less than15.762, 31.476 and 44.566 is 80% ,90% and 95% respectively. This is very important for the prediction of geomagnetic storm. In order to predict a geomagnetic storm and ap, we could find the probability of GIC that is no less than certain value from the table 3, thus helping us to avoid economic losses caused by underestimation or overestimation of GIC.

Table 3: Fitted values of different quantile

tau \\ ap	0.5	0.6	0.7	0.8	0.9	0.95
111	6.317	6.975	10.329	12.272	21.685	30.273
132	7.322	7.989	12.007	13.788	24.709	34.687
154	8.374	9.05	13.765	15.376	27.876	39.311
179	9.57	10.257	15.762	17.18	31.476	44.566
207	10.91	11.608	18	19.201	35.507	50.451
236	12.297	13.008	20.317	21.295	39.683	56.547
300	15.359	16.097	25.431	25.915	48.898	69.999

4. Conclusions

This article is trying to put forward a new method of power weighted quantile regression (PWRQ), especially when there are a lot of knots in statistics, PWQR is proved to be superior to EWQR. This article proves that all the weighting of those whose actual value is less than its estimated value by PWQR and all the weighting proportion are very close to the corresponding τ; that all the weighting of those whose actual value is more than its estimated value by PWQR and all the weighting proportion are very close to the corresponding $1 - \tau$. In this article, it also gives examples to testify the superiority of power weighted quantile regression to quantile regression. In the end, by power weighted quantile regression, we could get GIC value (which has little difference with actual value), and predict the probability of GIC in a certain range, reducing unnecessary losses.

Appendix: Proof of Theorem 1

Proof of Theorem 1: The EWQR objective function $R(\beta)$ is

$$R(\beta) = \min_{\beta \in R^p} \sum_{i=1}^{n} ||x_i||^{\alpha} \rho_{\tau}(y_i - x_i^T \beta)$$

$$= \min_{\beta \in R^p} \sum_{i=1}^{n} ||x_i||^{\alpha} \rho_{\tau}(y_i - x_i^T \beta)(\tau - I(y_i < x_i^T \beta))$$

The function is piecewise linear continuous function, which is not differentiable at the point $y_i = x_i^T \beta$. But there exists directional derivatives at all the points. In order to compute the minimum of $R(\beta)$, we consider its directional derivatives. The directional derivative of the objective function in direction w is given by

$$\nabla R(\beta, w) = \frac{d}{ds} R(\beta + sw)|_{s=0}$$

$$= \frac{d}{ds} \sum_{i=1}^{n} ||x_i||^{\alpha} (y_i - x_i^T \beta - x_i^T sw)[\tau - I(y_i < x_i^T \beta + x_i^T sw)]|_{s=0}$$

$$= -\sum_{i=1}^{n} ||x_i||^{\alpha} \psi_{\tau}(y_i - x_i^T \beta, -x_i^T w) x_i^T w$$

where

$$\psi_{\tau}(u, v) = \begin{cases} \tau - I(u < 0) \; if \; u \neq 0 \\ \tau - I(u < 0) \; if \; u = 0 \end{cases}$$

The parameter vector $\hat{\beta}$ minimizes of $R(\beta)$ if and only if $\nabla R(\beta, w) \geq 0$ for $\forall$ w.

$$R(\hat{\beta}, w) = -\sum_{i=1}^{n} ||x_i||^{\alpha} \psi_{\tau}(y_i - x_i^T \beta, -x_i^T w) x_i^T w$$

Because X includes an intercept term, there exists a vector $\gamma \in R^p$ such that $x_i^T \gamma = 1, i = 1, 2, \ldots, n$. If we let $w = -\gamma$ in expression (4), we get

$$\sum_{i=1}^{n} ||x_i||^{\alpha} \psi_{\tau}(y_i - x_i^T \hat{\beta}, 1) \geq 0$$

That is

$$\tau \sum_{i=1}^{n} ||x_i||^{\alpha} I(y_i > x_i^T \hat{\beta}) + (\tau - 1) \sum_{i=1}^{n} ||x_i||^{\alpha} I(y_i < x_i^T \hat{\beta}) + \tau \sum_{i=1}^{n} ||x_i||^{\alpha} I(y_i = x_i^T \hat{\beta}) \geq 0 \; (5)$$

The inequalities of expressions (5) can be rewritten as the inequalities of expression:

$$\tau[\sum_{i=1}^{n} ||x_i||^{\alpha} I(y_i > x_i^T \hat{\beta}) + \sum_{i=1}^{n} ||x_i||^{\alpha} I(y_i < x_i^T \hat{\beta})$$

$$+ \sum_{i=1}^{n} ||x_i||^{\alpha} I(y_i = x_i^T \hat{\beta})] \geq \sum_{i=1}^{n} ||x_i||^{\alpha} I(y_i < x_i^T \hat{\beta}) \qquad (6)$$

Form (6)

$$\tau \sum_{i=1}^{n} ||x_i||^{\alpha} \geq \sum_{i=1}^{n} ||x_i||^{\alpha} I(y_i < x_i^T \hat{\beta})$$

Thereby

$$\frac{\sum_{i=1}^{n} ||x_i||^{\alpha} I(y_i < x_i^T \hat{\beta})}{\sum_{i=1}^{n} ||x_i||^{\alpha}} \leq \tau$$

So the left of (2) has proved.

Both sides subtracting

$$-\sum_{i=1}^{n}||x_i||^{\alpha}\, I\!\left(y_i > x_i^T\hat{\beta}\right) - \sum_{i=1}^{n}||x_i||^{\alpha}\, I\!\left(y_i = x_i^T\hat{\beta}\right)$$

into (5),we obtain

$$1 - \tau \le \frac{\sum_{i=1}^{n}||x_i||^{\alpha}\, I\!\left(y_i > x_i^T\hat{\beta}\right) + \sum_{i=1}^{n}||x_i||^{\alpha}\, I\!\left(y_i = x_i^T\hat{\beta}\right)}{\sum_{i=1}^{n}||x_i||^{\alpha}}$$

So the right of (3) has proved.

In a similar way, we can let $w = \gamma$ (so than $x_i^T w = 1$), the right of (2) and the left of (3) are proved

Acknowledgments

References

[1] He, F. X. and Ma, X. J.(2011). A new method for assessment of GIC using geomagentic indices. *Multimedia Technology (ICMT), 2011 International Conference on* 1900-1902.

[2] Hendricks,W. and Koenker,R.(1991).Hierarchical spline models for conditional quantiles and the demand for electricity. *Journal of the American Statistical Association* **87**, 58-68.

[3] Taylor,J. W.(2008)Using exponentially weighted quantile regression to estimate value at risk and expected shortfall. *Journal of Financial Econometrics* **6**, 382-406.

[4] Koenker R.(2005). *Quantile regression*.Cambridge University Press, London.

[5] Trichtchenko, L. and Boteler,D.H.(2004). Modeling geomagnetically induced currents using geomagnetic indices and data. *IEEE Transactions on Plasma Science* **32**,1459-1467.

[6] Parzen, M. I., Wei,L.,Ying, Z..(1994). A resampling method based on pivotal estimating functions. *Biometrika* **81**, 341–350.

[7] Koenker. R. and Machado,J. A. F. (1999). Goodness of fit and related infer¬ence processes for quantile regression. *Journal of the American Statistical Association* **94**, 1296-1310.

Xue J. Ma
Center for Applied Statistics
Renmin University of China
59 Zhongguancun Ave., Beijing, 100872, China
yinuoyumi@ruc.edu .cn

Feng X. He
Department of Mathematics and Physics
North China Electric Power University
2 Beinong Road, Huilongguan Town, Changping District, Beijing,102206, China
he fx@126.com