Nongaussian time series model via Quantile Regression

Marcelo Ruas, *Student Member, IEEE,* and Alexandre Street, *Member, IEEE*

Abstract

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

Keywords

Quantile Regression, Model Identification, Non-gaussian time series model

TODO LIST

complete this 1
ct gas 2

I. INTRODUCTION

Renewable energy power is in the vogue in the recent years. This is happening because it is a much cleaner way of producing energy than by using other sources, such as coal and gas, and with less hazard potential as the nuclear plants. The installed capacity of renewable plants has been increasing in a fast pace. In spite of being a , it brings a new problem for the energy planners: due to its unpredictability behavior, the cost for the system may be high. Having better models can help the planner to make better and less risky choices to the system, increasing the renewable energy attractiveness.

complete this

The conventional way of having a model to produce point forecasts is not useful when dealing with renewable energy, as the variability and having notion of risk is of extreme importance for planning. Our attention is, then, turned to probabilistic forecasting models. Among these, we have many possibilities. [1] reviews the commonly used methodologies for this task, separating them in the parametric and nonparametric classes. Main characteristics of **parametric models** are assuming a distribution shape and low computational costs. GAS models and ARIMA-GARCH, for example, model the renewable series by knowing the distribution *a priori*. On the other hand, **nonparametric models** don't require a distribution to be specified, but needs mode data to produce a good approximation and have a higher computational cost. Popular methods are Quantile Regression, Kernel density estimation, Artificial Intelligence and a mix of them.

ct

gas

A common finding is that wind and solar time series don't have a Gaussian behavior. Furthermore, we can not point any distribution which fits on data without a doubt. Because of that, we choose to use a nonparametric approach. Quantile Regression (QR) is a powerful tool for measuring quantiles others than the median or predicting the mean. The approach of QR we use is the one defined in [2]. In this article, we use quantile regression and regularization techniques, such as the best subset selection [3] and the ℓ_1 penalization [4]. The AdaLasso variant, where each coefficient may have a different weight on the objective function to ensure oracle proprieties, is developed on [5].

The next section discusses with bigger details how to fit a distribution function $Q_{Y|X}(\alpha, x)$ from a sequence of estimated quantiles, as well as showing two different strategies to estimate them: linear models and nonparametric models. In the former, q_α is a linear function of an explanatory variable x_t. In the latter, we let $q_\alpha(x_t)$ assume any functional form. To prevent overfitting, however, we penalize the function's roughness by incorporating a penalty on the second derivative.

As opposed to the conventional work of doing a model to forecast the conditional mean, our work focus on finding a distribution for y_t on each t.

We find a time series model, based on quantile autoregression (as in Koenker 2005).

As we are interested in the whole distribution of $\hat{y}_{t+k|t}$, we estimate a phin grid of quantiles in $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{|A|} < 1$, such that the distribution can be well approximated.

As a Quantile Autoregression model, we are interested in selecting the best subset of variables do model the time series.

As we are trying to model the whole k-step ahead distribution, we estimate many quantiles. We didn't find any previous work where a given α-quantile model influenced another model.

In all works found, each quantile is estimated separately.

Regularization can be done by introducing a penalty on the ℓ_1-norm of the coefficients. The work by [4] defines proprieties and convergence analysis. The AdaLasso variant, where each coefficient may have a different weight on the objective function to ensure oracle proprieties, is developed on [5].

II. Quantile Regression based time series model

Let Y be a real valued random variable. The quantile function $Q_Y : [0,1] \to \mathbb{R}$ is defined pointwise by its α-quantile, which is given by

$$Q_Y(\alpha) = F_Y^{-1}(\alpha) = \inf\{y : F_Y(y) \geq \alpha\}, \tag{1}$$

where F_Y is the distribution function of random variable Y and $\alpha \in [0,1]$. Equation 1 defines what we call from now on as the quantile function $Q_Y(\cdot)$, in relation to random variable Y. In this article, we are interested in a conditional quantile function $Q_{Y|X=x} : [0,1] \times \mathbb{R}^d \to \mathbb{R}$ (in short, from now on, $Q_{Y|X}(\cdot,\cdot)$), where X can be a vector.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. The conditional quantile function can be found as the result of the following optimization problem:

$$\begin{aligned} Q_{Y|X}(\alpha,\cdot) \quad \in \quad \underset{q_\alpha(\cdot)}{\arg\min} \quad & (1-\alpha)\int_{\omega\in\Omega} |Y(w) - q_\alpha(X(w))|^- P(dw) \qquad (2) \\ & +(\alpha)\int_{\omega\in\Omega} |Y(w) - q_\alpha(X(w))|^+ P(dw) \\ & q_\alpha \in \mathcal{Q}, \qquad (3) \end{aligned}$$

The argument of optimization problem described on equation 2 is the function q_α, which belongs to a function space $\mathcal{Q}$. We might have different assumptions for the space $\mathcal{Q}$, depending on the type of function we want to find for q_α. A few properties, however, must be achieved by our choice. The conditional quantile function $Q_{Y|X}(\alpha)$ must be monotone on α, and its first derivative must be limited.

In this work, we use the sample quantile, where we calculate the optimization based on a finite number of observations, instead of integrating over all domain of random variable Y. For

the specific case where the random variable is a time series y_t, quantiles are estimated from a n size sample of observations of y_t and a explanatory variable x_t for each t, such that our random sample is formed by the sequence $\{y_t, x_t\}_{t=1}^{n}$. To estimate the α-quantile from a sample, we change 2 for the following optimization problem:

$$\hat{Q}_{Y|X}(\alpha, \cdot) \quad \in \quad \underset{q_\alpha(\cdot)}{\arg\min} \sum_{t \in T} \alpha |y_t - q_\alpha(\cdot)|^+ + \sum_{t \in T} (1-\alpha)|y_t - q_\alpha(\cdot)|^-, \tag{4}$$

$$q_\alpha \in \mathcal{Q}, \tag{5}$$

where $T = \{1, \ldots, n\}$, $|x|^+ = \max\{0, x\}$ and $|x|^- = -\min\{0, x\}$. The solution from the above problem is an estimator $\hat{Q}_{Y|X}$ for the quantile function $Q_{Y|X}$.

To model this problem as a Linear Programming problem, thus being able to use a modern solver to fit our model, we create variables ε_t^+ e ε_t^- to represent $|y - q(\cdot)|^+$ and $|y - q(\cdot)|^-$, respectively. The optimal argument $q_\alpha^*(\cdot)$ on the Linear Programming problem 6 is the estimated α-quantile for the given random sample.

$$\begin{aligned} q_\alpha^*(.) \in \underset{q_\alpha(\cdot), \varepsilon_t^+, \varepsilon_t^-}{\arg\min} & \sum_{t \in T} \left(\alpha \varepsilon_t^+ + (1-\alpha)\varepsilon_t^-\right) \\ \text{s.t. } & \varepsilon_t^+ - \varepsilon_t^- = y_t - q_\alpha(x_t), \qquad \forall t \in T, \\ & \varepsilon_t^+, \varepsilon_t^- \geq 0, \qquad \forall t \in T. \end{aligned} \tag{6}$$

Let A be a set containing a sequence of probabilities α_i such that $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_Q < 1$. This set represents a finite discretization of the interval $[0, 1]$. One of our goals with quantile regression is to estimate a quantile function $\hat{Q}_{Y|X}$ of a given real valued random variable X from a sequence of quantiles $q_{\alpha_1}(x_t) \leq q_{\alpha_2}(x_t) \leq \cdots \leq q_{\alpha_{|A|}}(x_t)$, with $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_{|A|} < 1$, for any given t. The process of fitting $\hat{Q}_{Y|X}$ is by mapping every α_i with its estimated quantile $\hat{q}_{\alpha_i}(x_t)$. The denser the grid of values in A, better is the approximation of $Q_{Y|X}$. Thus, the distribution found for Y is nonparametric, as no previous assumptions are made about its shape, and its form is fully recovered by the data we have.

A typical problem, however, arises when working with quantile regression. When quantiles are estimated independently, it is possible to find $q_\alpha(x_t) > q_{\alpha'}(x_t)$, for a given t, when $\alpha_1 < \alpha_2$. An example can be seen on Figure **??**, where quantiles $\alpha = 0.95$ and $\alpha = 0.9$ cross. This problem, called *crossing quantiles*, can be prevented by estimating all quantiles with a single minimization problem.

As opposed to traditional parametric models where one has to assume the error distribution on the model

$$y_t = h(F_t) + \varepsilon_t, \tag{7}$$

as we choose the nonparametric methodology we are released from this step. A distribution for ε_t is defined within the estimation procedure. The predictive distribution is derived from the individual quantiles

$$\hat{q}^{\alpha}_{t+k|k} = \hat{y}_{t+k|t} + \hat{F}^{\varepsilon}_{t,k}(\alpha). \tag{8}$$

[6] defines a model based on extreme machine learning to estimate every α-quantile in a thin grid $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_Q < 1$.

A. Linear parametric explanatory model

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

B. General non-linear nonparametric model

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut porttitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.

REFERENCES

[1] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, Apr. 2014. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032114000446

[2] R. Koenker, *Quantile regression*. Cambridge university press, 2005, no. 38.

[3] D. Bertsimas, A. King, and R. Mazumder, "Best Subset Selection via a Modern Optimization Lens," *arXiv:1507.03133 [math, stat]*, Jul. 2015, arXiv: 1507.03133. [Online]. Available: http://arxiv.org/abs/1507.03133

[4] A. Belloni and V. Chernozhukov, "L1-Penalized Quantile Regression in High-Dimensional Sparse Models," *arXiv:0904.2931 [math, stat]*, Apr. 2009, arXiv: 0904.2931. [Online]. Available: http://arxiv.org/abs/0904.2931

[5] G. Ciuperca, "Adaptive LASSO model selection in a multiphase quantile regression," *Statistics*, vol. 50, no. 5, pp. 1100–1131, Sep. 2016. [Online]. Available: http://dx.doi.org/10.1080/02331888.2016.1151427

[6] C. Wan, J. Lin, J. Wang, Y. Song, and Z. Y. Dong, "Direct Quantile Regression for Nonparametric Probabilistic Forecasting of Wind Power Generation," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2767–2778, Jul. 2017.