PONTÍFICIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

Quantile Regression

Marcelo Castiel Ruas*, Henrique Helfer Hoeltgebaum†, Alexandre Street‡, Cristiano Fernandes§

January 16, 2017

*Aluno de doutorado do Departamento de Engenharia Elétrica da PUC-RIO.
†Aluno de doutorado do Departamento de Engenharia Elétrica da PUC-RIO.
‡Professor do Departamento de Engenharia Elétrica da PUC-RIO.
§Professor do Departamento de Engenharia Elétrica da PUC-RIO.

1 Simulation

In this section, we investigate how to simulate future paths of the time series y_t. Let T be the total number of observations of y_t. We produce S different paths with size K for each. We have T observations of y_t and we want to simulate Given a vector of explanatory variables x_t, let q_t^α be given by the following linear model:

$$q_t^\alpha = \beta_0^\alpha + x_t^T \beta^\alpha + \varepsilon_t, \tag{1.1}$$

where β^α is a vector of coefficients for the explanatory variables. The variables chosen to compose x_t can be either exogenous variables, autoregressive components of y_t or both. As the distribution of ε_t is unknown, we have to use a nonparametric approach in order to estimate its one-step ahead density.

The coefficients β_0^α and β^α are the solution of the minimization problem given in equation **??**, reproduced here for convenience:

$$\begin{aligned} \min_{\beta_0,\beta,\varepsilon_t^+,\varepsilon_t^-} & \sum_{t=1}^{n} \left(\alpha\varepsilon_t^+ + (1-\alpha)\varepsilon_t^-\right) \\ \text{s.t. } & \varepsilon_t^+ - \varepsilon_t^- = y_t - \beta_0 - \beta^T x_t, \qquad \forall t \in \{1,\dots,n\}, \\ & \varepsilon_t^+, \varepsilon_t^- \geq 0, \qquad \forall t \in \{1,\dots,n\}. \end{aligned} \tag{1.2}$$

To produce S different paths of $\{\hat{y}_t\}_{t=T+1}^{T+K}$, we use the following procedure:

1. For every quantile $\alpha_i \in (0,1)$, we use equation 1.1 to produce a forecast of $\hat{q}_{T+1}^{\alpha_i}$, as x_{T+1} is supposed to be known at time $T+1$. In the presence of exogenous variables that are unknown, it is advisable to incorporate its uncertainty by considering different scenarios. In each scenario, though, x_{T+1} must be considered fully known.

2. In any given t, by choosing many different values of α_i, we can estimate a sequence of quantiles $q_t^{\alpha_1} \leq q_t^{\alpha_2} \leq \cdots \leq q_t^{\alpha_Q}$ with $0 < \alpha_1 < \alpha_2 < \cdots < \alpha_Q < 1$. Let $F_{y_{T+1}}$ be the estimated distribution function of y_{T+1}. The process of fitting $\hat{F}_{y_{T+1}}$ is by mapping every α_i with its estimated quantile $\hat{q}^{\alpha_i}$. A problem arises for the distribution extremities, because when $\alpha = 0$ or $\alpha = 1$, the problem 1.2 becomes unbounded. In order to find good estimates for y_{T+1} when $F_{y_{T+1}}$ approaches 0 or 1, we can either use a kernel smoothing function, splines, linear approximation, or any other method. **This will be developed later.** When this sequence of chosen α_i is thin enough, we can approximate well the distribution function of y_{T+1}, as is shown in Figure . Thus, the distribution found for $\hat{y}_{T+1}$ is nonparametric, as no previous assumptions are made about its shape, and its form is fully recovered by the data we have.

3. Once we have a distribution for y_{T+1}, we can generate K different simulated values, drawn from the distribution $\hat{F}_{y_{T+1}}$ found on step 2.

Let X be a random variable with uniform distribution over the interval $[0,1]$. By using results from the Probability Integral Transform, we know that the random variable $F_{y_{T+1}}^{-1}(X)$ has the same distribution as y_{T+1}. So, by drawing a sample of size K from X and applying the inverse function of $F_{y_{T+1}}$, we have our sample of size K for y_{T+1}.

References

[1] Dimitris Bertsimas, Angela King, and Rahul Mazumder. Best subset selection via a modern optimization lens. *arXiv preprint arXiv:1507.03133*, 2015.

[2] Seung-Jean Kim, Kwangmoo Koh, Stephen Boyd, and Dimitry Gorinevsky. ℓ_1 trend filtering. *SIAM review*, 51(2):339–360, 2009.

[3] Roger Koenker. *Quantile regression*. Number 38. Cambridge university press, 2005.

[4] Fernando Porrua, Bernardo Bezerra, Luiz Augusto Barroso, Priscila Lino, Francisco Ralston, and Mario Pereira. Wind power insertion through energy auctions in brazil. In *Power and Energy Society General Meeting, 2010 IEEE*, pages 1–8. IEEE, 2010.