

# Regressão Quantílica

---

Henrique Helfer Hoeltgebaum\*, Marcelo Castiel Ruas†, Alexandre Street‡,  
Cristiano Fernandes§  
8 de abril de 2016

## 1 Introdução

A previsão da maioria dos modelos de regressão é apenas uma estimativa da média condicional da variável resposta, para um dado conjunto de preditores. Porém, a média condicional representa apenas a medida de tendência central da distribuição da variável resposta. Um quadro mais completo a respeito da relação entre a distribuição condicional da variável dependente com a das independentes pode ser obtido pelos seus quantis. Recentes avanços em computação tornaram viáveis o desenvolvimento de modelos de regressão para prever um dado quantil da distribuição condicional, tanto parametricamente quanto não parametricamente. Tal metodologia é chamada de Regressão Quantílica, mas a metodologia (da estimação do quantil condicional) se aplica para qualquer modelo estatístico.

## 2 Regressão Quantílica

Em um modelo de regressão simples via mínimos quadrados ordinários (OLS), modelamos a média condicional da variável resposta (ou dependente) como função de uma ou mais variáveis independentes. Mas, como a média não é uma descrição completa da distribuição, modelar a média não resulta em uma descrição completa da relação entre a

---

\*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.

variável dependente e as respectivas variáveis independentes, logo essa análise pode não ser adequada.

Para enfatizar a necessidade de tal análise, cita-se um simples exemplo, supondo que nossa variável dependente seja bimodal ou multimodal, isto é, possui múltiplos picos conforme ilustrado pela Figura 2.1.

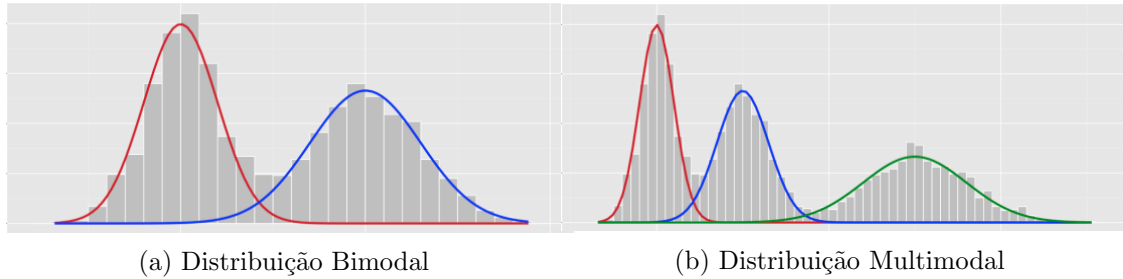


Figura 2.1: Gráficos das densidades de distribuições bimodais e multimodais

Se soubermos o que causa a bimodalidade, podemos separar aquela variável e prosseguir com uma análise estratificada, porém caso não conhecermos, o arcabouço dos modelos de regressão quantílica se demonstram uma alternativa eficaz para tal análise. Uma vez que a regressão por mínimos quadrados, nesse caso pode causar resultados não confiáveis dado que esta se apoia na média como uma medida central para uma distribuição bimodal.

Além do exemplo supracitado, a regressão quantílica traz um resumo de informações muito mais completo, com relações interessantes que podem ser obtidas através dos quantis da distribuição condicional da variável resposta. Quando existe o interesse em saber, por exemplo, relações entre as pessoas com maior e menor renda, conforme será apresentado na Seção 2.1, cada parâmetro dos regressores são estimados como função do quantil a ser estudado, nesse caso estudaríamos os quantis  $\{0.05, 0.95\}$  da distribuição da variável dependente *renda* com suas covariáveis, tendo para cada quantil  $\tau$  um respectivo  $\hat{\beta}_i(\tau)$ .

Ou seja, ao invés do modelo de regressão linear estimado via mínimos quadrados ordinários,

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_{in} x_{in} + \varepsilon_i \quad (2.1)$$

teremos o seguinte modelo de regressão quantílica, que fornece os quantis condicionais de  $y$  por meio da seguinte equação,

$$Q_y(\tau | x_1, x_2, \dots, x_n) = \beta_0(\tau) + \beta_1(\tau)x_1 + \beta_2(\tau)x_2 + \dots + \beta_n(\tau)x_n + F_\varepsilon^{-1}(\tau) \quad (2.2)$$

onde  $F_\varepsilon$  indica a função distribuição dos erros.

## 2.1 Estimação dos parâmetros

Conforme apresentado na Seção 1.3 de [3], os pontos discutidos de que os quantis podem ser expressos como solução de um simples problema de otimização leva, naturalmente, a métodos mais gerais de estimação de modelos de quantil condicional. Mínimos

quadrados oferecem o referido arcabouço para esse desenvolvimento. Sabendo que a média amostral resolve o problema

$$\min_{\mu \in \mathbb{R}} \sum_{i=1}^n (y_i - \mu)^2 \quad (2.3)$$

sugerindo que, se desejamos expressar a média condicional de  $y$  dado  $x$  como  $\mu(x) = x^T \beta$ , então  $\beta$  pode ser estimado resolvendo

$$\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n (y_i - x_i^T \beta)^2. \quad (2.4)$$

Similarmente, uma vez que o  $\tau$ -ésimo quantil amostral,  $\hat{\alpha}(\tau)$  resolve

$$\min_{\alpha \in \mathbb{R}} \sum_{i=1}^n \rho_{\tau}(y_i - \alpha). \quad (2.5)$$

somos levados a especificar o  $\tau$ -ésimo quantil condicional como  $Q_y(\tau|x) = x^T \beta(\tau)$ , e para a estimação de  $\hat{\beta}(\tau)$  basta resolver

$$\min_{(\beta, u, v) \in \mathbb{R}^p} \sum_{i=1}^n \rho_{\tau}(y_i - x_i^T \beta). \quad (2.6)$$

Esta é a idéia principal apresentada com maiores detalhes em [4].

O problema de regressão quantílica apresentado pela Equação (2.6) pode ser reformulado como um problema de programação linear,

$$\min_{\beta \in \mathbb{R}^p \times \mathbb{R}_+^{2n}} \{ \tau \mathbf{1}_n^T u + (1 - \tau) \mathbf{1}_n^T v \mid X\beta + u - v = y \}, \quad (2.7)$$

onde  $X$  representa a matriz usual dos regressores  $n$  por  $p$ . Divide-se o vetor de resíduos  $y - X\beta$  em partes negativas e positivas ( $u$  e  $v$ ), e então estamos minimizando uma função linear em um poliedro restrito por um conjunto de restrições, e muitas das propriedades importantes das soluções,  $\hat{\beta}(\tau)$ , os quais são denominamos "quantis regressores", são derivadas imediatamente a partir de propriedades de soluções bem conhecidas de problemas lineares.

## 2.2 Seleção do modelo

Conforme apresentado na Seção 4.9.1 de [3], existe uma vasta literatura a respeito de seleção de modelos. Serão abordados apenas alguns desenvolvimentos recentes a cerca do tema.

Em [6] é apresentada uma variação do critério de penalização de Schwarz para M-estimadores de regressão, incluindo o caso da regressão quantílica. É enfatizada a necessidade de procedimentos invariantes a locação e escala: para regressões da mediana ( $\tau=1/2$ ) o seguinte cálculo rende o critério,

$$SIC(j) = \log(\hat{\sigma}_j) + \frac{1}{2} p_j \log n, \quad (2.8)$$

onde  $\hat{\sigma}_j = \sum_{i=1}^n \rho_{1/2}(y_i - x_i^T \hat{\beta}_n(1/2))$  e  $p_j$  indica a dimensão do  $j$ -ésimo modelo. Esse procedimento rende uma seleção de modelo consistente sob a circunstância de que um dos modelos  $p_j$ -dimensional esteja de fato bem especificado.

Alternativamente ao critério supracitado, podemos considerar o critério de Akaike, calculado conforme

$$AIC(j) = \log(\hat{\sigma}_j) + p_j. \quad (2.9)$$

Com probabilidade positiva, a seleção baseada na minimização do  $AIC(j)$  superestima a dimensão do modelo, mas pode vir a exercer uma performance superior para a previsão.

## 2.3 Diagnósticos

Em [5] é apresentada uma versão do familiar  $R^2$  do modelo de regressão por mínimos quadrados para o modelo de regressão quantílica, o  $R^1(\tau)$ . Atenta-se ao fato de que esse novo parâmetro é uma medida de diagnóstico para um quantil em particular, ao contrário do seu análogo  $R^2$  que fornece uma medida global de ajuste sob toda a distribuição condicional. É fácil de imaginar circunstâncias na qual as covariáveis podem exercer um efeito em uma cauda da distribuição condicional da variável resposta porém podem vir a não surtir efeito na outra cauda. Em [1] é apresentado um exemplo disso no arcabouço de regressão quantílica.

Alternativamente à metodologia supracitada, em [7] (Seção 5), é apresentada uma ferramenta de diagnóstico visual para verificar o ajuste do modelo aos dados. Essa metodologia é denominada de "gráfico da minhoca", nela é possível apontar os quantis nos quais o ajuste pode ser melhorado e comparar o ajuste de diferentes modelos com o modelo selecionado.

## 2.4 Previsão fora da amostra

Escrever sobre previsão pontual e densidade preditiva

# 3 Regressão Quantílica para séries temporais

Quando relaxada a condição de independência das observações, podemos continuar com a aplicação da metodologia de regressão quantílica. Bloomfield e Steiger (1983) mostraram a normalidade assintótica de um estimador autoregressivo da mediana para um modelo no qual as observações  $\{(x_i, y_i) : i = 1, \dots, n\}$  eram assumidas serem estacionárias e de diferença martingal ergódica.

## 3.1 Autoregressão

Já existe uma ampla literatura a respeito da estimação e inferência para modelos de séries temporais lineares se utilizando de métodos de regressão quantílica. Bloomfield e Steiger (1983) consideraram o modelo autoregressivo

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \dots + \alpha_p y_{t-p} + u_t \quad (3.1)$$

com a condição usual de estacionariedade de que as raízes do polinômio característico  $1 - \alpha_1 z - \dots - \alpha_p z^p = 0$  permaneçam fora do círculo unitário e com erros  $u_t$  iid com uma distribuição nonlattice  $F$  satisfazendo as seguintes condições: (i)  $E(\log_+ |u_t|)$ , (ii) mediana  $\{u_t\}=0$ , e (iii)  $E|u_t| < \infty$ . Isso prova que o estimador da regressão da mediana

$$\hat{\alpha}_n = \operatorname{argmin}_{\alpha \in \mathbb{R}} \sum_{t=1}^n |y_t - \alpha_0 - \alpha_1 y_{t-1} - \dots - \alpha_p y_{t-p}| \quad (3.2)$$

É fortemente consistente. Quando  $F$  tem densidade  $f(0) > 0$ , e  $E u_t^2 = \sigma^2 < \infty$ , as técnicas descritas anteriormente (veja Poolard, 1991, Exemplo 2) podem ser usadas para estabelecer normalidade assintótica de  $\hat{\alpha}_n(\tau)$ .

É valido considerar um simples modelo AR(1) a fim de ilustração,

$$y_t = \alpha_1 y_{t-1} + u_t, \quad (3.3)$$

demonstraremos algumas características especiais da especificação AR. Considerando  $n^{-1} \sum y_{t-1}^2 \rightarrow \frac{\sigma^2}{(1-\alpha^2)}$  e utilizando a quantidade que podemos denominar de  $D_0$  para enfatizar a analogia com o caso de regressão, é fácil mostrar que

$$\sqrt{n}(\hat{\alpha}_n - \alpha) \rightsquigarrow N(0, \omega^2 D_0^{-1}) \quad (3.4)$$

onde  $\omega = (2f(0))^{-1}$ . Em contraste, o estimador de mínimos quadrados de  $\alpha$ ,

$$\tilde{\alpha}_n = \frac{\sum_{t=2}^n y_t y_{t-1}}{\sum_{t=2}^n y_{t-1}^2} \quad (3.5)$$

satisfaz, sob a condição de que,

$$\sqrt{n}(\tilde{\alpha}_n - \alpha) \rightsquigarrow N(0, \sigma^2 D_0^{-1}) \quad (3.6)$$

## 3.2 Modelos ARMA

Segundo a Seção 2.6.2 de [3] os modelos autoregressivos se ajustam bem ao arcabouço original da regressão quantílica: a computação é direta e muitas das propriedades da teoria assintótica de regressão se mantêm.

No caso dos modelos ARMA, a situação é complicada pela não convexidade da função objetivo, uma dificuldade que é superada pela aproximação criteriosa da função objetivo na vizinhança do verdadeiro valor do parâmetro. Isso assegura uma sequência de minimizadores *locais* que convergem como especificados mas também implica em alguma preocupação a respeito dos métodos computacionais que irão levar ao minimizador apropriado.

Até hoje, grande parte da literatura de aplicações de regressão quantílica em séries temporais foca em modelos com inovações iid. Parece haver um escopo considerável para explorar modelos que partem dessa condição restritiva. Os primeiros passos nessa direção são descritos na Seção 8.3 de [3], através da exemplificação dos modelos QAR que será discutido na próxima Seção 3.3.

### 3.3 Modelos QAR

Existe uma ampla literatura a respeito dos modelos QAR. Nesse modelo o  $\tau$ -ésimo quantil condicional da variável resposta  $y_t$  é expressado como uma função linear de valores defasados da resposta. Mas uma característica notável a respeito dessa literatura é que ela tenha se focado, quase que exclusivamente em inovações iid, na qual as variáveis condicionantes possuem um papel de deslocar a locação da densidade condicional de  $y_t$ , porém não surtindo efeito na escala nem na forma condicional.

Será explorado o método de estimação e inferência em uma classe mais geral de modelos QAR onde todos os coeficientes autoregressivos assumem ser  $\tau$  dependentes e, portanto, são capazes de alterar a locação, escala e forma da densidade condicional. Escreveremos a forma geral do modelo como

$$Q_{y_t}(\tau|y_{t-1}, \dots, y_{t-p}) = \beta_0(\tau) + \beta_1(\tau)y_{t-1} + \dots + \beta_p(\tau)y_{t-p} \quad (3.7)$$

ou matricialmente podemos escrever como,

$$Q_{y_t}(\tau|I_{t-1}) = x_t^T \alpha(\tau) \quad (3.8)$$

onde  $I_t$  representa o conjunto de informações geradas por  $\{y_s, s \leq t\}$ , e o vetor de repostas defasadas é representado por  $x_t = (1, y_{t-1}, y_{t-2}, \dots, y_{t-p})^T$ .

Vamos motivar o uso deste modelo pela utilização do mesmo em sua forma mais simples, o QAR(1) representado por

$$Q_{y_t}(\tau|I_{t-1}) = \alpha_0(\tau) + \alpha_1(\tau)y_{t-1}. \quad (3.9)$$

Sabendo que  $Q_y(\tau|I_{t-1})$  é o quantil condicional da densidade de  $y_t$ , o modelo pode ser representado por

$$y_t = \alpha_0(U_t) + \alpha_1(U_t)y_{t-1} \quad (3.10)$$

onde assume-se que  $U_t$  seja uma variável aleatória uniformemente distribuída no intervalo unitário e iid. Simular o modelo é convenientemente conduzido empregando essa forma. O modelo AR(1) gaussiano clássico é obtido ao atribuirmos  $\alpha_0(u) = \sigma\Phi^{-1}(u)$  e  $\alpha_1$ , uma constante. A formulação pela equação (3.10) revela que o modelo pode ser interpretado como uma versão não usual de modelos autoregressivos de coeficientes aleatórios. Porém, em contraste com a maior parte da literatura em modelos de coeficientes aleatórios - veja, Nicholls e Quinn (1982) e Tjostheim (1986) - no qual os coeficientes são assumidos serem estocasticamente independentes de cada um, o modelo QAR possui coeficientes que são funcionalmente dependentes. Uma vez que a monotonicidade é necessária pela função quantil, veremos que isso implica em uma certa disciplina na forma assumida pela função- $\alpha$ . Essa disciplina essencialmente requer que o vetor  $\alpha(\tau)$ , ou alguma transformação afim deste, seja monotônica em cada coordenada. Essa condição assegura que o vetor aleatório  $\alpha_t = \alpha(U_t)$  seja comonotônico.

A estimação do modelo linear QAR envolve a resolução do problema

$$\min_{\alpha \in \mathbb{R}^{p+1}} \sum_{t=1}^n \rho_\tau(y_t - x_t^T \alpha). \quad (3.11)$$

Soluções  $\hat{\alpha}(\tau)$  podem ser chamadas quantis autoregressivos. Dado  $\hat{\alpha}(\tau)$ , o  $\tau$ -ésimo quantil condicional de  $y_t$ , condicionado ao conjunto de informações passadas, pode ser estimado por

$$\hat{Q}_{y_t}(\tau|x_{t-1}) = x_t^T \hat{\alpha}(\tau) \quad (3.12)$$

e a densidade condicional de  $y_t$  pode ser estimada pela diferença entre os quocientes,

$$\hat{f}_{y_t}(\tau|x_{t-1}) = (\tau_i - \tau_{i-1}) / (\hat{Q}_{y_t}(\tau_i|x_{t-1}) - \hat{Q}_{y_t}(\tau_{i-1}|x_{t-1})) \quad (3.13)$$

para alguma sequência apropriadamente escolhidas de  $\tau$ s.

Quantiles	(Intercept)	series[, 2]
0.05	7.660 (-11.634, 66.602)	0.639 (-0.640, 0.999)
0.25	37.623 ( 22.217, 39.891)	0.169 ( 0.107, 0.466)
0.75	60.886 ( 17.305, 69.763)	-0.139 (-0.305, 0.725)
0.95	81.008 ( 49.484, 92.594)	-0.401 (-0.551, 2.151)

Quantiles	(Intercept)	series[, 2]
0.05	7.660 (-11.634, 66.602)	0.639 (-0.640, 0.999)
0.25	37.623 ( 22.217, 39.891)	0.169 ( 0.107, 0.466)
0.75	60.886 ( 17.305, 69.763)	-0.139 (-0.305, 0.725)
0.95	81.008 ( 49.484, 92.594)	-0.401 (-0.551, 2.151)

Nosso objetivo principal neste trabalho é estimar a função  $\mathcal{Q}_{y_t|y_{t-p}}^\alpha(t)$  que corresponde ao quantil  $\alpha$  quando previsto  $p$  passos  $\tilde{A}$  frente, para uma dada série temporal  $y_t$ , como a apresentada na figura 3.1. Então, dada uma sequência  $\{y_t\}$ , podemos fazer uma correspondência entre cada observação  $y_t$  e sua defasagem de ordem  $p$ . A figura 3.2 mostra um exemplo dessa relação.

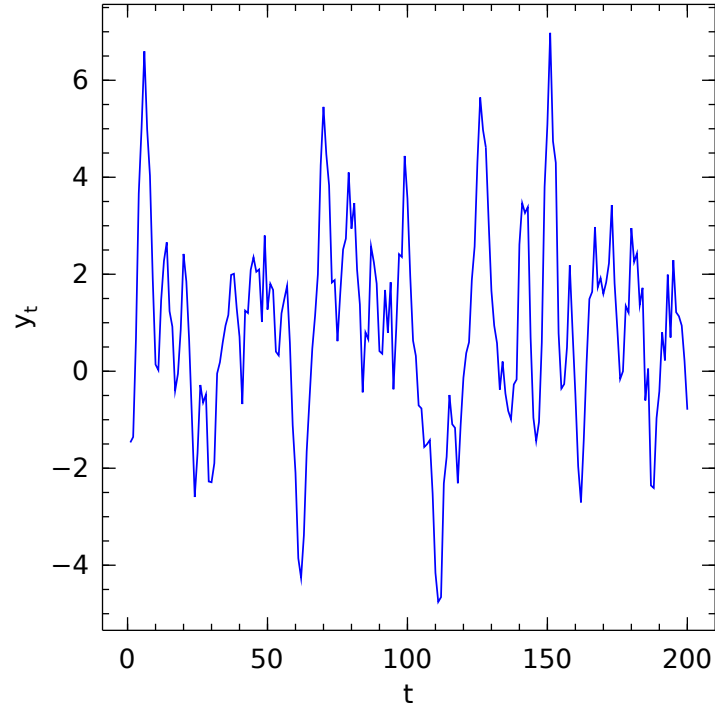


Figura 3.1: Série temporal  $y_t$

Refazer gráfico com pontos maiores

Vamos investigar duas maneiras de estimar os quantis da relação acima: utilizando um modelo linear e um modelo não paramétrico.

## 4 Estimando a QAR com um modelo linear

Ao estimar um modelo de regressão quantílica, estamos interessados em encontrar uma função  $f$  que minimize a soma dos erros

$$\min_f \sum_{i=1}^n \alpha |y_t - f(t)|^+ + (1 - \alpha) |y_t - f(t)|^-, \quad (4.1)$$

em que  $|x|^+ = \max\{0, x\}$  e  $|x|^- = -\min\{0, x\}$ . Para resolver este problema como um problema de programação linear, podemos criar variáveis  $\delta^+$  e  $\delta^-$  para representar as



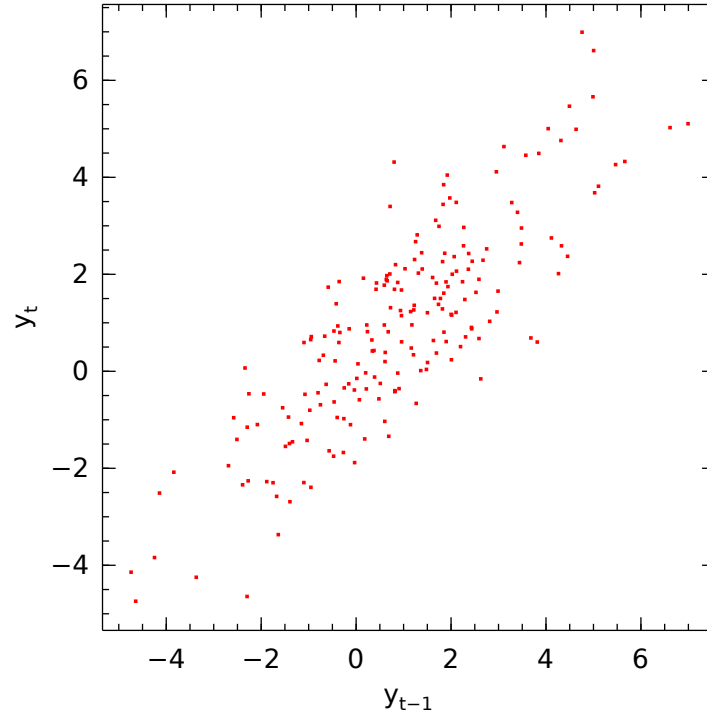


Figura 3.2: Relationship between  $y_t$  and its first lag  $y_{t-1}$

funções  $|\cdot|^+$  e  $|\cdot|^-$ , respectivamente, e montá-lo da seguinte forma:

$$\begin{aligned} \min_f \sum_{i=1}^n (\lambda \delta_i^+ + (1 - \lambda) \delta_i^-) \\ \text{sujeito à } \delta_i^+ - \delta_i^- = y_i - f(x_i), \quad \forall i \in \{1, \dots, n\}, \\ \delta_i^+, \delta_i^- \geq 0, \quad \forall \{1, \dots, n\}. \end{aligned} \quad (4.2)$$

Quando assumimos que a função  $f$  que estima o  $\alpha$ -ésimo quantil é uma função linear, tal que  $f(x) = \beta_0 + \beta^T x$ , os valores de  $\beta_0$  e  $\beta$  podem ser obtidos como sendo aqueles vindos da solução do seguinte problema de minimização:

$$\begin{aligned} \min_{\beta_0, \beta} \sum_{i=1}^n (\lambda \delta_i^+ + (1 - \lambda) \delta_i^-) \\ \text{sujeito à } \delta_i^+ - \delta_i^- = y_i - \beta_0 - \beta^T x_i, \quad \forall i \in \{1, \dots, n\}, \\ \delta_i^+, \delta_i^- \geq 0, \quad \forall \{1, \dots, n\}. \end{aligned} \quad (4.3)$$

Resultados para a estimação linear

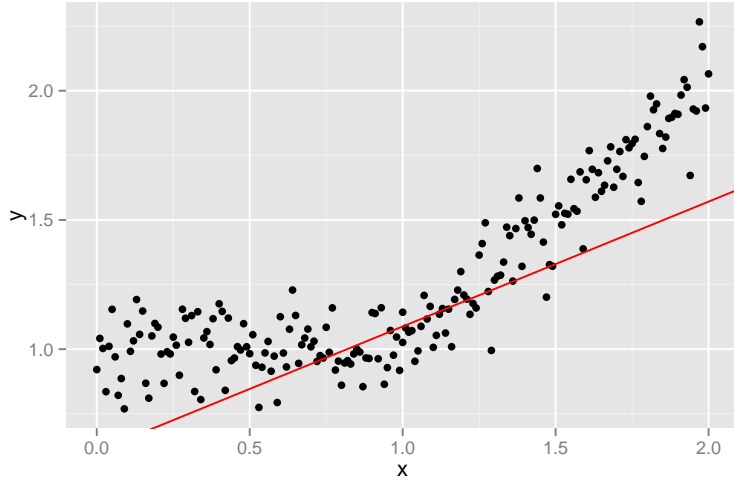


Figura 5.1: Example of data where nonlinearity is present and a linear quantile estimator is employed

## 5 Estimando a QAR não parametricamente

Fitting a linear estimator for the Quantile Auto Regression isn't appropriate when nonlinearity is present in the data. This nonlinearity may produce a linear estimator that underestimates the quantile for a chunk of data while overestimating for the other chunk (we illustrate this in figure 5.1). To prevent this issue from occurring we propose a modification which we let the prediction  $\mathcal{Q}_{y_t|y_{t-1}}^\alpha(t)$  adjust freely to the data and its nonlinearities. To prevent overfitting and smoothen our predictor, we include a penalty on its roughness by including the  $\ell_1$  norm of its second derivative. For more information on the  $\ell_1$  norm acting as a filter, one can refer to [2].

Let  $\{\tilde{y}_t\}_{t=1}^n$  be the sequence of observations in time  $t$ . Now, let  $\tilde{x}_t$  be the  $p$ -lagged time series of  $\tilde{y}_t$ , such that  $\tilde{x}_t = L^p(\tilde{y}_t)$ , where  $L$  is the lag operator. Matching each observation  $\tilde{y}_t$  with its  $p$ -lagged correspondent  $\tilde{x}_t$  will produce  $n - p$  pairs  $\{(\tilde{y}_t, \tilde{x}_t)\}_{t=p+1}^n$  (note that the first  $p$  observations of  $y_t$  must be discarded). When we order the observation of  $x$  in such way that they are in growing order

$$\tilde{x}_{(p+1)} \leq \tilde{x}_{(p+2)} \leq \dots \leq \tilde{x}_{(n)},$$

we can then define  $\{x_i\}_{i=1}^{n-p} = \{\tilde{x}_{(t)}\}_{t=p+1}^n$  and  $\{y_i\}_{i=1}^{n-p} = \{\tilde{y}_{(t)}\}_{t=p+1}^n$  and  $I = \{2, \dots, n - p - 1\}$ . As we need the second difference of  $q_i$ ,  $I$  has to be shortened by two elements.

Our optimization model to estimate the nonparametric quantile is as follows:

$$\begin{aligned} \mathcal{Q}_{y_t|y_{t-1}}^\alpha(i) = \arg \min_{q_i} \sum_{i \in I} (|y_i - q_i|^+ \alpha + |y_i - q_i|^- (1 - \alpha)) \\ + \lambda \sum_{i \in I} |D^2 q_i|, \end{aligned} \quad (5.1)$$

where  $D^2q_t$  is the second derivative of the  $q_t$  function, calculated as follows:

$$D^2q_i = \left( \frac{q_{i+1} - q_i}{x_{i+1} - x_i} \right) - \left( \frac{q_i - q_{i-1}}{x_i - x_{i-1}} \right).$$

The first part on the objective function is the usual quantile regression condition for  $\{q_i\}$ . The second part is the  $\ell_1$ -filter. The purpose of a filter is to control the amount of variation for our estimator  $q_i$ . When no penalty is employed we would always get  $q_i = y_i$ . On the other hand, when  $\lambda \rightarrow \infty$ , our estimator approaches the linear quantile regression.

The output of our optimization problem is a sequence of ordered points  $\{(x_i, q_i)\}_{i \in I}$ . The next step is to interpolate these points in order to provide an estimation for any other value of  $x$ . To address this issue, we propose a B-splines interpolation, that will be discussed in another subsection.

When estimating quantiles for a few different values of  $\alpha$ , however, sometimes we find them overlapping each other, which we call crossing quantiles. To prevent this, we include a non-crossing constraint:

$$q_i^\alpha \leq q_i^{\alpha'}, \quad \forall i \in I, \alpha < \alpha'. \quad (5.2)$$

traduzir trecho acima

Testamos a regressão quantílica não paramétrica para estimar quantis para diversos tipos de dados e de penalizações. A seguir, apresentamos o resultado da regressão quantílica para dados gerados a partir de um modelo ARIMA com diversos valores distintos de penalização, com ou sem a restrição de não-cruzamento de quantis. As figuras contém todos os quantis entre 1% e 99%.

Na estimação com valor de  $\lambda = 0.3$  (figura 5.2), observamos que a estimação do quantil possui uma variância muito alta. Nesses casos, o erro *in-sample* é baixo, mas as previsões serão muito ruins. À medida que o valor incrementamos o valor de  $\lambda$  (figura 5.4, onde utilizamos  $\lambda = 1$ ), a função estimada para os quantis fica cada vez mais próxima de uma função linear, até que para valores mais extremos de *lambda* (ver figura 5.5, onde utilizamos  $\lambda = 10$ ) ela será totalmente linear.

A restrição 5.2, ao forçar que a função de nenhum quantil se cruze, produz uma diferença nos valores estimados. Esta diferença, no entanto, é perceptível apenas nas extremidades dos gráficos. As figuras e exemplificam as diferenças da inclusão ou não desta restrição.

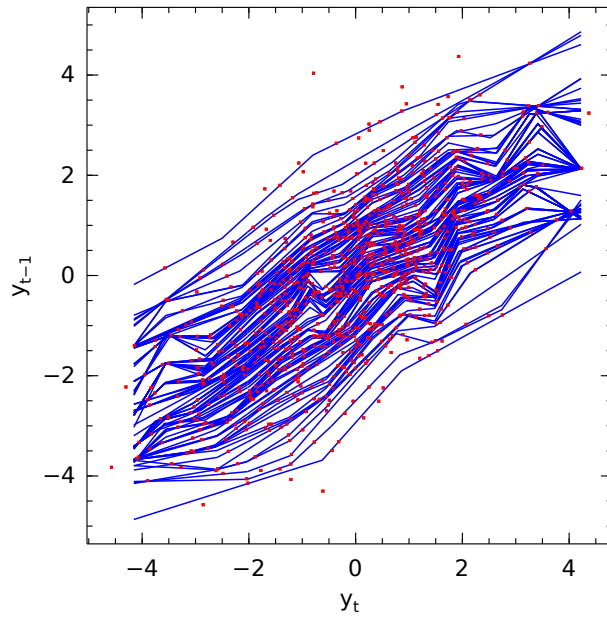


Figura 5.2: Dados simulados ARIMA com penalização  $\lambda = 0.3$

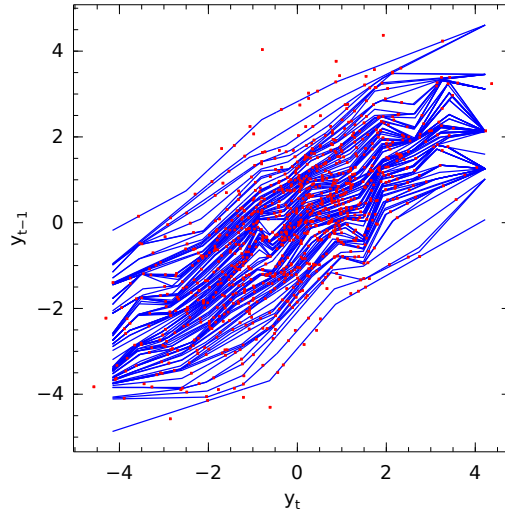


Figura 5.3: Dados simulados ARIMA com penalização  $\lambda = 0.3$  e restrição de não cruzamento de quantis

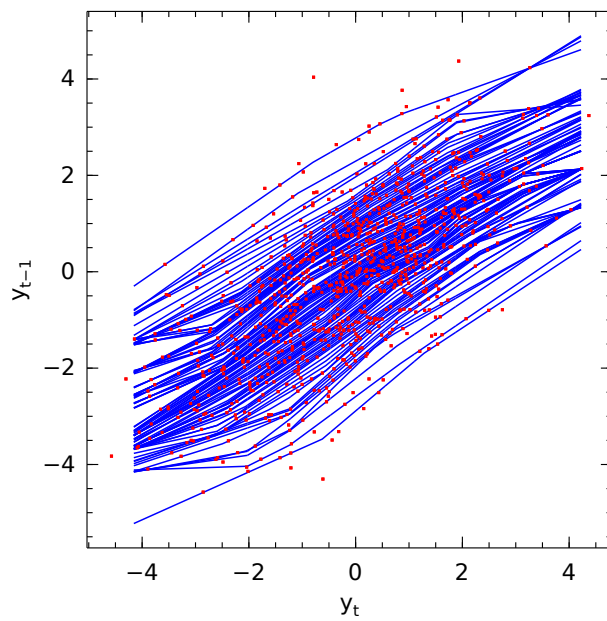


Figura 5.4: Dados simulados ARIMA com penalização  $\lambda = 1$

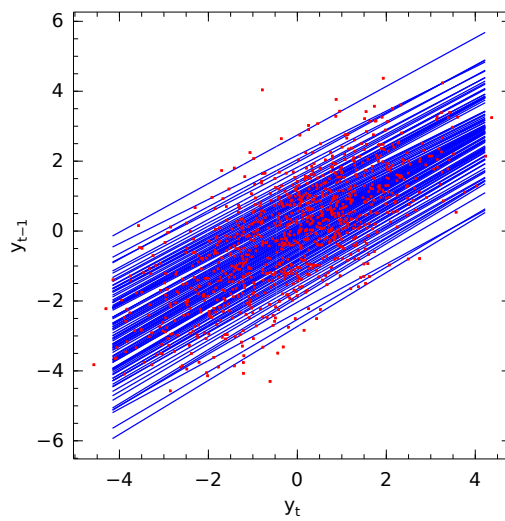


Figura 5.5: Dados simulados ARIMA com penalização  $\lambda = 10$

## Referências

- [1] Gary Chamberlain. Quantile regression, censoring, and the structure of wages.
- [2] Seung-Jean Kim, Kwangmoo Koh, Stephen Boyd, and Dmitry Gorinevsky.  $\ell_1$  trend filtering. *SIAM review*, 51(2):339–360, 2009.
- [3] Roger Koenker. *Quantile regression*. Number 38. Cambridge university press, 2005.
- [4] Roger Koenker and Gilbert Bassett Jr. Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50, 1978.
- [5] Roger Koenker and Jose AF Machado. Goodness of fit and related inference processes for quantile regression. *Journal of the american statistical association*, 94(448):1296–1310, 1999.
- [6] Jose AF Machado. Robust model selection and m-estimation. *Econometric Theory*, 9:478–493, 1993.
- [7] Stef van Buuren. Worm plot to diagnose fit in quantile regression. *Statistical Modelling*, 7(4):363–376, 2007.