Ficha de Leitura

Marcelo Ruas

November 23, 2016

Quantile Regression

Takeuchi 2006

Define the loss function as

$$I_\tau(\xi) = \begin{cases} \tau\xi & \text{if } \xi \geq 0 \\ (\tau - 1)\xi & \text{if } \xi < 0 \end{cases}$$

Expected quantile risk is defined as

$$R[f] = E_{\rho(x,y)}[I_\tau(y - f(x))].$$

Since $\rho(x, y)$ is unknown, they employ a regularizer:

$$R_{reg}[f] = \frac{1}{m}\sum_{i=1}^{m}[I_\tau(y - f(x_i))] + \frac{\lambda}{2}\|g\|_H^2, \text{ where } f = g + b \text{ and } b \in \mathbb{R}.$$

Here $\|\cdot\|_H$ is the Reproducing Kernel Hilbert Space and the constant offset b is not regularized.

- Reproducing Kernel Hilbert Space: Let X be an arbitrary set and H a Hilbert space of real-valued functions on X. The evaluation functional over the Hilbert space of functions H is a linear functional that evaluates each function at a point x ,

  $$L_x : f \mapsto f(x) \forall f \in H.$$

  We say that H is a reproducing kernel Hilbert space if, for all $x \in X$, L_x is continuous at any $f \in H$ or, equivalently, if L_x is a bounded operator on H , i.e. there exists some $M > 0$ such that

  $$|L_x[f]| := |f(x)| \leq M\|f\|_H \forall f \in H.$$

To ensure the *non-crossing constraints*, so that $0 \leq \tau_1 \leq \tau_2 \leq \cdots \leq \tau_n \leq 1$ at any given point $\{x_j\}_{j=1}^{l}$, they add aditional constraints. The function f for the τ_h-th quantile is defined as a basis expansion $f_h(x) = \langle\phi(x), w_h\rangle + b_h$ for $h = 1, 2, \ldots, n$. In H, the non-crossing constraints are represented as linear constraints

$$\langle\phi(x_j), w_h\rangle + b_h \leq \langle\phi(x_j), w_{h+1}\rangle + b_{h+1}, \quad \text{for all } 1 \leq h \leq n-1,\ 1 \leq j \leq l.$$

It is worth noting that, after enforcing the non-crossing constraints, the quantile property as in Lemma 3 may not be guaranteed. This is because the method both tries to optimize for the quantile property and the non-crossing property (in relation to other quantiles). Hence, the final outcome may not empirically satisfy the quantile property. Yet, the non-crossing constraints are very nice because they ensure the semantics of the quantile definition: lower quantile level should not cross the higher quantile level.

The performance is checked with respect to two criteria:

- Expected risk with respect to the l_τ loss function.
- Ensure that we produce numbers $f_\tau(x)$ which exceed y with probabilitu close to τ. The quantile property is measured by *ramp loss*.

Variable Selection for Nonparametric Quantile Regression via Smoothing Spline AN OVA ()

Abstract: *We tackle the problem via regularization in the context of smoothing spline ANOVA models. The proposed sparse nonparametric quantile regression (SNQR) can identify important variables and provide flexible estimates for quantiles.*

Variable selection in quantile regression is much more difficult than that in the least squares regression. The variable selection is carried at various levels of quantiles, which amounts to identifying variables that are important for the entire distribution, rather than limited to the mean function as in the least squares regression case. This has important applications to handle heteroscedastic data.

A multivariate function $f(x) = f(x^{(j)}, \ldots, x^{(d)})$ has the ANOVA decomposition:

$$f(x) = b + \sum_{j=1}^{d} f_j\left(x(j)\right) + \sum_{j<k} f_{j,k}\left(x^{(j)}, x^{(k)}\right) + \ldots$$

where b is a constant, f_j's are the main effects and $f_{j,k}$'s are the two-way interactions, and so on.