SPATIAL QUANTILE REGRESSION WITH SMOOTH DENSITY FUNCTIONS

Halley Brantley, Montserrat Fuentes, Joseph Guinness, and Eben Thoma

NC State University, Virginia Commonwealth University, Cornell University

U.S. Environmental Protection Agency

Supplementary Material

S1 Differentiability at τ_U

Let Y have a quantile function as defined in (2.1) and (2.2) with an I-spline basis order greater than $q+1$ and a density that is continuous and $(q-1)^{th}$ order differentiable at $Q(\tau_U)$. If α_U is constrained so that Eq. S1.1 does not result in $\theta_{M-q,p} < 0$, then Y has a density that is q^{th}-order differentiable at $Q(\tau_U)$ for any $\mathbf{x} \in \mathbb{R}_+^P$ if and only if

$$\theta_{M-q,p} = \frac{1}{I_{M-q}^{(q+1)}(\tau_U)} \left\{ \frac{\sigma_{U,p}}{\alpha_U(\tau_U - 1)^{q+1}} (-\alpha_U - q)_{q+1} - \sum_{m=M-q+1}^{M} \theta_{m,p} I_m^{(q)}(\tau_U) \right\} \tag{S1.1}$$

where $I_{M-q}^{(q+1)}(\tau_U)$ is the $(q+1)^{th}$ order derivative of the $(M-q)^{th}$ I-spline basis function, $(-\alpha_U - q)_{q+1} = \prod_{j=0}^{q}(-\alpha_U - j)$.

S2 Proofs

Proof of Proposition 1

Proof. We will only prove the case for the lower tail, the proof for the upper tail is equivalent. Given the assumptions we have already shown $Q'(\tau) > 0$ for all τ. Thus the density exists and is given by $f(y) = \frac{1}{Q'(Q^{-1}(y))}$. The density is continuous at τ_L if and only if $Q'(\tau)$ is continuous at τ_L. By definition only a single I-spline basis function has a non-zero derivative at τ_L: $I_1'(\tau_L)$. Therefore, the following is a necessary and sufficient condition for a continuous density for any value of $x_p \geq 0$:

$$Q'(\tau_L) = \sum_{p=1}^{P} \sigma_{L,p} x_p \tau_L^{-1} = \sum_{m=1}^{M} \sum_{p=1}^{P} \theta_{m,p} x_p I_m'(\tau_L) = \sum_{p=1}^{P} \theta_{1,p} x_p I_1'(\tau_L) \quad \text{(S2.1)}$$

Now since $x_p \geq 0$ is defined arbitrarily, take $x_p = 0$ for all $p \neq 1$, that is $x_p = 0$ for all predictors other than the intercept term, $x_1 = 1$. Then (S2.1) is equivalent to

$$\theta_{1,1} = \frac{\sigma_{L,1}}{\tau_L * I_1'(\tau_L)}. \quad \text{(S2.2)}$$

Similarly, for $p > 1$ take $x_p \neq 0$ for some q and $x_p = 0$ for all $p > 1$ and $p \neq q$. Then (S2.1) and (S2.2) are equivalent to

$$\theta_{1,q} = \frac{\sigma_{L,q}}{\tau_L * I_1'(\tau_L)}. \tag{S2.3}$$

Hence we have proved Proposition 1. □

Proof of Theorem 1

Proof. Let Y have a quantile functions as defined in Eq. 2.1 (main text) with an I-spline basis order greater than $q+1$ and a density that is $(q-1)^{th}$ order differentiable. Given $\tau \leq \tau_L$,

$$\beta_p(\tau) = \theta_{1,p} - \frac{\sigma_{L,p}}{\alpha_L}\left[\left(\frac{\tau}{\tau_L}\right)^{-\alpha_L} - 1\right]$$

and $Q(\tau) = \sum_{p=1}^{p} x_p \beta_p(\tau)$. The density of Y is q^{th} order differentiable if and only if $Q(\tau)$ is $(q+1)^{th}$ order differentiable. By definition $Q(\tau)$ is $(q+1)^{th}$ order differentiable at all points except τ_L and τ_U. $Q(\tau)$ is $(q+1)^{th}$ order differentiable at τ_L if and only if,

$$\sum_{p=1}^{P} x_p \sum_{m=0}^{M} \theta_{m,p} I_m^{(q+1)}(\tau_L) = \sum_{p=1}^{p} x_p \beta_p^{(q+1)}(\tau_L) \tag{S2.4}$$

$I_m^{(q+1)}(\tau_L) = 0$ for $m = 0$ and $m > q+1$ so eq. S2.4 is equivalent to

$$\sum_{p=1}^{P} x_p \sum_{m=1}^{q+1} \theta_{m,p} I_m^{(q+1)}(\tau_L) = \sum_{p=1}^{p} x_p \beta_p^{(q+1)}(\tau_L) \tag{S2.5}$$

Because $\beta(\tau)$ is a polynomial in τ we can write

$$\beta_p^{(q+1)}(\tau) = \frac{-\sigma_{L,p}\tau_L^{\alpha_L}}{\alpha_L} \tau^{-\alpha_L - q - 1} \prod_{j=0}^{q} (-\alpha_L - j) \tag{S2.6}$$

$$\beta_p^{(q+1)}(\tau_L) = \frac{-\sigma_{L,p}}{\alpha_L \tau_L^{q+1}} (-\alpha_L - q)_{q+1} \tag{S2.7}$$

Now since $x_p \geq 0$ is defined arbitrarily, take $x_p = 0$ for all $p \neq 1$, that is $x_p = 0$ for all predictors other than the intercept term, $x_1 = 1$. We further start with the case with $q = 1$. By proposition 1, $\theta_{1,p}$ can be written as a function of $\sigma_{L,p}$ and eq. S2.5 is satisfied if and only if

$$\theta_{1,1} I_1^{(2)}(\tau_L) + \theta_{2,1} I_2^{(2)}(\tau_L) = \sigma_{L,1} \tau_L^{-2}(-\alpha_L - 1) \tag{S2.8}$$

$$\theta_{2,1} = \frac{1}{I_2^{(2)}(\tau_L)} \left[\sigma_{L,1} \tau_L^{-2}(-\alpha_L - 1) - \theta_{1,1} I_1^{(2)}(\tau_L) \right] \tag{S2.9}$$

Similarly, for $p > 1$ take $x_p \neq 0$ for some q and $x_p = 0$ for all $p > 1$ and $p \neq q$. Then (S2.5) and (S2.8) are equivalent to

$$\theta_{2,p} = \frac{1}{I_2^{(2)}(\tau_L)} \left[\sigma_{L,p} \tau_L^{-2}(-\alpha_L - 1) - \theta_{1,p} I_1^{(2)}(\tau_L) \right] \tag{S2.10}$$

We have thus shown that we can ensure 1^{st} order differentiability of the density function of Y at $Q(\tau_L)$ by constraining $\theta_{1,p}$ and $\theta_{2,p}$ as functions of $\sigma_{L,p}$ and α_L for all p. Returning to the more general case, given a density of Y that is $(q-1)^{th}$ order differentiable, we again take $x_p = 0$ for all $p \neq 1$. Then the density of Y is q^{th} order differentiable if and only if

$$\sum_{m=1}^{q+1} \theta_{m,1} I_m^{(q+1)}(\tau_L) = \frac{-\sigma_{L,1}}{\alpha_L \tau_L^{q+1}} (-\alpha_L - q)_{q+1} \tag{S2.11}$$

$$\theta_{q+1,1} = \frac{1}{I_{q+1}^{(q+1)}(\tau_L)} \left[\frac{-\sigma_{L,1}}{\alpha_L \tau_L^{q+1}} (-\alpha_L - q)_{q+1} - \sum_{m=1}^{q} \theta_{m,1} I_m^{(q+1)}(\tau_L) \right] \tag{S2.12}$$

Similarly, for $p > 1$ take $x_p \neq 0$ for some q and $x_p = 0$ for all $p > 1$ and $p \neq q$. Then (S2.5) and (S2.11) are equivalent to

$$\theta_{q+1,p} = \frac{1}{I_{q+1}^{(q+1)}(\tau_L)} \left[\frac{-\sigma_{L,p}}{\alpha_L \tau_L^{q+1}} (-\alpha_L - q)_{q+1} - \sum_{m=1}^{q} \theta_{m,p} I_m^{(q+1)}(\tau_L) \right] \tag{S2.13}$$

We have thus proved Theorem 1. $\square$

S3 Expectation and Covariance

$$
\begin{aligned}
&E[Y_t(s)|\Theta(s), X_t(s)] \\
=&\int_0^1 Q_Y[\tau|\Theta(s), X_t(s)]d\tau \\
=&\sum_m \sum_p \theta_{m,p}(s)x_{p,t}(s)G_m \\
+&\sum_p \left(\tau_L\theta_{1,p}(s)x_{p,t}(s) + (1-\tau_U)x_{p,t}(s)\sum_m \theta_{m,p}(s) + \frac{\sigma_{L,p}(s)x_{p,t}(s)\tau_L}{\alpha_L(s)-1} + \frac{(1-\tau_U)\sigma_{U,p}(s)x_{p,t}(s)}{1-\alpha_U(s)} \right)
\end{aligned}
$$

where $G_m = \int_{\tau_L}^{\tau_U} I_m(\tau)d\tau$. As the last two terms approach zero, the distribution of the marginal expectation of Y becomes a linear combination of the $\theta_{m,p}$ which have log normal distributions.

The marginal expectation of $Y(s)$, marginalizing over $\theta_{m,p}$ and σ is

$$
\begin{aligned}
&E[Y_t(s)|X_t(s)] \\
=&\sum_m \sum_p \mu_{m,p}x_{p,t}(s)G_m \\
+&\sum_p \left(\tau_L\mu_{1,p}x_{p,t}(s) + (1-\tau_U)x_{p,t}(s)\sum_m \mu_{m,p} \right) + \\
+&\sum_p \left(\tau_L^2\mu_{2,p}I_2'(\tau_L)x_{p,t}(s)E\left[\frac{1}{\alpha_L-1}\right] + (1-\tau_U)^2\mu_{M,p}I_M'(\tau_U)x_{p,t}(s)E\left[\frac{1}{1-\alpha_U}\right] \right)
\end{aligned}
$$

Based on the model structure, given Θ, $Y_t(s)$ and $Y_t(s')$ are independent. Thus the conditional covariance is zero and

$$
E[Y_t(s)Y_t(s')|X_t(s)\Theta(s)] = E[Y_t(s)|X_t(s)\Theta(s)]E[Y_t(s')|X_t(s')\Theta(s')]
$$

$$\begin{aligned}
E[Y_t(s)Y_t(s')|X_t(s)] =& E_\Theta\Bigg[E\big[Y_t(s)|X_t(s),\Theta(s)\big]\,E\big[Y_t(s')|X_t(s'),\Theta(s')\big]\Bigg] \\
=& \sum_m G_m^2 x_{p,t}(s)x_{p,t}(s')[\Sigma_{m,p}(s,s') + \mu_{m,p}^2] \\
&+ \sum_m \sum_p \{x_{p,t}(s)G_m\mu_{m,p}(s)\} \sum_{(l,k)\neq(m,p)} \{x_{k,t}(s')G_l\mu_{l,k}\}
\end{aligned}$$

Bibliography

E. Hansen and M. Patrick. A family of root finding methods. *Numerische Mathematik*, 27(3):257–269, 1976.