

Quantile estimation of Correlated Spatio-Temporal Process

January 16, 2022

1 Introduction

Quantile regression is an useful method to quantify the effect of a covariate on the distribution of response. The usual regression on the conditional mean of the response variable may be inadequate in certain applications. For example, in weather data, it is very useful to analyze the occurrence of extreme heat (or extreme cold) events, rather than the average temperature conditions.

2 Proposed Estimator

We consider a spatio-temporal setup similar to the one described in [Deb and Jana \[2021\]](#). Let Y_{tj} to be the observation at t -th timepoint for the j -th location space, for $t = 1, 2, \dots, n$ and $j = 1, 2, \dots, p$. Also, assume that there exists a time-dependent covariate $X_t \in \mathbb{R}^d$, which controls the general trend across all locations.

$$\mathbf{Y}_t = \begin{pmatrix} Y_{t1} \\ Y_{t2} \\ \vdots \\ Y_{tp} \end{pmatrix} = \mu(X_t) + \Sigma^{1/2}(X_t) \begin{pmatrix} e_{it} \\ e_{it} \\ \vdots \\ e_{tp} \end{pmatrix} \quad (2.1)$$

where $\mu(\cdot) : \mathbb{R}^d \rightarrow \mathbb{R}^p$ and $\Sigma(X_t)$ is a $m \times m$ positive definite matrix depending on the covariates X_t , and e_{tj} 's are errors which are independent and identically distributed. We will make the following assumption about the covariates.

- (A1) There exists a sequence of random variables $\{\epsilon_t\}_{t=-\infty}^{\infty}$ with natural filtration \mathcal{F}_t independent of the random variables e_{tj} s such that X_t is $\mathcal{F}_{(t-1)}$ measurable for any $t = 0, \pm 1, \pm 2, \dots$. That implies $X_t = h(\dots, \epsilon_{-1}, \epsilon_0, \epsilon_1, \dots, \epsilon_{(t-1)})$ for a measurable function h .

With such a setup as given in Eq. (2.1), the quantity of interest would be the quantiles of the p -variate distribution of \mathbf{Y}_t . There have been numerous notions of quantiles for multivariate distribution using convex hull [[Eddy, 1985](#)], minimum volume sets [[Einmahl and Mason, 1992](#)], geometric generalization of spatial median [[Chaudhuri, 1996](#)], center outward quantile functions [[Figalli, 2018](#)] etc. Among this, the geometric notion of quantile developed by [Chaudhuri \[1996\]](#) is a direct generalization of the usual quantile regression framework by [Koenker and Bassett Jr \[1978\]](#), and hence it also fits nicely into the spatio-temporal framework in our consideration. He defines the multivariate quantile for a d -variate random vector \mathbf{X} as

$$Q(\mathbf{u}) = \arg \min_{\mathbf{q} \in \mathbb{R}^d} \mathbb{E} [\Phi(\mathbf{u}, \mathbf{X} - \mathbf{q})],$$

for any $u \in B^d = \{x \in \mathbb{R}^d : \|x\| < 1\}$, where $\Phi(\mathbf{u}, \mathbf{t}) = \|\mathbf{t}\| + \langle \mathbf{u}, \mathbf{t} \rangle$. This notion generalizes the idea of [Koenker and Bassett Jr \[1978\]](#) who viewed the quantile estimation problem as an optimization problem of the pinball loss function. Note that, the choice of $\mathbf{u} \in B^d$ towards the boundary leads to more extreme quantiles, while the choice of \mathbf{u} towards the origin leads to the spatial centers. Bahadur's representation and various asymptotic analysis of this geometric notion of multivariate quantiles is given by [Chaudhuri \[1996\]](#).

In the spatio-temporal setup as in Eq. (2.1), if we consider only a single location (i.e., $p = 1$) and assume the model as only a temporal model, then a natural nonparametric quantile estimator would be

$$\arg \min_q \sum_{t=1}^n k(x, x_t) [|Y_t - q| + u(Y_t - q)], \quad (2.2)$$

where $k(\cdot, \cdot)$ is a kernel of appropriate bandwidth. For $u = (2\alpha - 1)$ with $0 < \alpha < 1$, this estimator becomes the kernel weighted version of the α -th quantile of the conditional distribution of Y_t conditioned on $x_t = x$. Therefore, a natural choice of estimator in the original spatio-temporal setup (with $p > 1$) is to introduce matrix valued kernel functions $K(\cdot, \cdot)$ (see [Alvarez et al. \[2011\]](#) for details) in the inner products. Thus, we consider the estimator

$$\hat{Q}(x, \mathbf{u}) = \arg \min_{\mathbf{q} \in \mathbb{R}^d} \left[\sum_{t=1}^n \|\mathbf{Y}_t - \mathbf{q}\|_{K^2(x, X_t)} + \langle \mathbf{u}, \mathbf{Y}_t - \mathbf{q} \rangle_{K(x, X_t)} \right], \quad (2.3)$$

where $\|\mathbf{Y}_t - \mathbf{q}\|_{K^2(x, X_t)} = \sqrt{(\mathbf{Y}_t - \mathbf{q})^\top K^2(x, X_t)(\mathbf{Y}_t - \mathbf{q})}$ and $\langle \mathbf{u}, \mathbf{Y}_t - \mathbf{q} \rangle_{K(x, X_t)} = \mathbf{u}^\top K(x, X_t)(\mathbf{Y}_t - \mathbf{q})$. The kernel $K(x, X_t)$ is normalized in the sense that it also contains a division by b_n where b_n is an appropriate normalizing constant depending on the bandwidth and sample size n . The population version of this estimator which we are trying to obtain is the conditional quantile of Y conditioned on the value of the temporal covariate $X = x$, which is given by

$$Q(x, \mathbf{u}) = \arg \min_{\mathbf{q} \in \mathbb{R}^d} \mathbb{E}_{\mathbf{Y}|\mathbf{X}=x} [\|\mathbf{Y} - \mathbf{q}\| + \langle \mathbf{u}, \mathbf{Y} - \mathbf{q} \rangle]. \quad (2.4)$$

3 Theoretical Properties

In this section, we wish to show that the estimator $\hat{Q}_n(x, \mathbf{u})$ is consistent for its population version $Q(x, \mathbf{u})$ for any choice of $\mathbf{u} \in \mathcal{U}$ as the number of timepoints $n \rightarrow \infty$, where

$$\mathcal{U} = \{\mathbf{u} : \mathbf{u}^\top \mathbf{u} \leq 1\}$$

We shall also try to investigate its asymptotic non-degenerate distribution of a scaled difference $\hat{Q}(x, \mathbf{u}) - Q(x, \mathbf{u})$. Let us also define the following notations for fixed \mathbf{u} and fixed value of x ,

$$\begin{aligned} \mathcal{L}_n(q) &= \sum_{i=1}^n [\|\mathbf{Y}_i - q\|_{K^2(x, X_i)} + \langle \mathbf{u}, \mathbf{Y}_i - q \rangle_{K(x, X_i)}] \\ \mathcal{L}(q) &= \mathbb{E} [\|Y(x) - q\| + \langle \mathbf{u}, Y(x) - q \rangle] \end{aligned}$$

where $Y(x)$ denotes that the random variable Y is a stochastic function of x , with a deterministic trend part and an additive random noise component.

We will start a result similar to the Theorem 2.1.2 of [Chaudhuri \[1996\]](#), which defines the estimating equation of the estimator.

Lemma 1. For any $x \in \mathbb{R}^d$ and any $\mathbf{u} \in \mathcal{U}$, the estimator $\widehat{Q}(x, \mathbf{u})$ satisfy

$$\sum_{t=1}^n K(x, X_t) \left[\frac{K(x, X_t)(\mathbf{Y}_t - \widehat{Q})}{\|K(x, X_t)(\mathbf{Y}_t - \widehat{Q})\|} + \mathbf{u} \right] = 0. \quad (3.1)$$

The true quantile $Q(x, \mathbf{u})$ satisfy the population version of the similar equation

$$\mathbb{E}_{\mathbf{Y}|\mathbf{X}=x} \left[\frac{Y - Q(x, \mathbf{u})}{\|Y - Q(x, \mathbf{u})\|} + \mathbf{u} \right] = 0. \quad (3.2)$$

Proof.

Note that, since $K_n(x, X_t)$ is a kernel matrix, it is positive definite. Therefore, the function $\Phi_K(\mathbf{u}, \mathbf{Y}_t - q) = \|\mathbf{Y}_t - q\|_{K^2(x, X_t)} + \langle \mathbf{u}, \mathbf{Y}_t - q \rangle_{K(x, X_t)}$ is convex in q . This implies that for any $h \in \mathbb{R}^d$, we must have

$$\begin{aligned} & \lim_{\epsilon \rightarrow 0+} \frac{1}{\epsilon} \sum_{i=1}^n \Phi_K(\mathbf{u}, \mathbf{Y}_t - \widehat{Q} + \epsilon h) - \Phi_K(\mathbf{u}, \mathbf{Y}_t - \widehat{Q}) \geq 0 \\ \Rightarrow & \sum_{t=1}^n \lim_{\epsilon \rightarrow 0+} \frac{\|\mathbf{Y}_t - \widehat{Q} + \epsilon h\|_{K^2(x, X_t)} - \|\mathbf{Y}_t - \widehat{Q}\|_{K^2(x, X_t)}}{\epsilon} + \lim_{\epsilon \rightarrow 0+} \frac{\langle \mathbf{u}, \epsilon h \rangle_{K(x, X_t)}}{\epsilon} \geq 0 \\ \Rightarrow & \sum_{t=1}^n \left[\frac{\langle \mathbf{Y}_t - \widehat{Q}, h \rangle_{K^2(x, X_t)}}{\|\mathbf{Y}_t - \widehat{Q}\|_{K^2(x, X_t)}} + \langle \mathbf{u}, h \rangle_{K(x, X_t)} \right] \geq 0, \quad \text{If } \widehat{Q} \neq \mathbf{Y}_t \text{ for any } t = 1, 2, \dots, n \end{aligned}$$

Since this holds for any $h \in \mathbb{R}^d$, we can replace $(-h)$ instead of h and we obtain the equality in Eq. (3.1). Eq. (3.2) follows similarly from Eq. (2.4). \square

Next, we shall present a generic result.

Lemma 2. Let, $\mathbf{X} \sim \mathcal{N}_p(\mu, \Sigma)$. Then

$$\begin{aligned} \mathbb{E} \left(\frac{\mathbf{X}}{\|\mathbf{X}\|} \right) &= \frac{\Gamma((p-1)/2)}{\sqrt{2}\Gamma(p/2)} \frac{(p-1)}{p} \Sigma^{1/2} \mu + o(p^{-3/2}) \\ \mathbb{E} \left(\frac{\mathbf{X}\mathbf{X}^\top}{\|\mathbf{X}\|^2} \right) &= \frac{1}{p} (\mathbf{I} + \Sigma^{1/2} \mu \mu^\top \Sigma^{1/2}) + \frac{\Sigma^{1/2} \mu \mu^\top \Sigma^{1/2}}{p(p-2)} + \\ & \quad \frac{(\text{trace}(\Sigma^{1/2}) - p + 1)(\text{trace}(\Sigma^{1/2}) - p + 3)}{p} \mathbf{I} - 2 \frac{(\text{trace}(\Sigma^{1/2}) - p + 2)}{p} \Sigma^{1/2} + \frac{1}{p} \Sigma + o(p^{-2}) \end{aligned}$$

Proof. Since, $\mathbf{X} \sim \mathcal{N}_p(\mu, \Sigma)$ we can rewrite $\mathbf{X} = \mu + \Sigma^{1/2} \mathbf{Z}$ where $\mathbf{Z} \sim \mathcal{N}_p(0, \mathbf{I})$. Let us rewrite $\frac{\mathbf{X}}{\|\mathbf{X}\|}$ as a function of μ and $\Sigma^{1/2}$, namely $h(\mu, \Sigma^{1/2}) = \frac{(\mu + \Sigma^{1/2} \mathbf{Z})}{\|\mu + \Sigma^{1/2} \mathbf{Z}\|}$. Then, an application of Taylor's theorem on h around the point $(0, \mathbf{I})$ yields that

$$\begin{aligned} & \mathbb{E} \left[\frac{\mathbf{X}}{\|\mathbf{X}\|} \right] \\ = & \mathbb{E} \left[\frac{\mathbf{Z}}{\|\mathbf{Z}\|} \right] + \mathbb{E} \left(\frac{\mathbf{I}}{\|\mathbf{Z}\|} - \frac{\mathbf{Z}\mathbf{Z}^\top}{\|\mathbf{Z}\|^3} \right) \Sigma^{1/2} \mu + \mathbb{E} \left(\left\langle \frac{\mathbf{I}}{\|\mathbf{Z}\|} - \frac{\mathbf{Z}\mathbf{Z}^\top}{\|\mathbf{Z}\|^3}, \Sigma^{1/2} - \mathbf{I} \right\rangle \mathbf{Z} \right) + o(\mathbb{E}(\|\mathbf{Z}\|^{-3})) \end{aligned}$$

$$\begin{aligned}
&= \frac{\Gamma((p-1)/2)}{\sqrt{2}\Gamma(p/2)} \frac{(p-1)}{p} \Sigma^{1/2} \mu + (\text{trace}(\Sigma^{1/2}) - p + 1) \mathbb{E} \left(\frac{\mathbf{Z}}{\|\mathbf{Z}\|} \right) - \mathbb{E} \left(\frac{(\mathbf{Z}^\top \Sigma^{1/2} \mathbf{Z}) \mathbf{Z}}{\|\mathbf{Z}\|^3} \right) + o(p^{-3/2}) \\
&= \frac{\Gamma((p-1)/2)}{\sqrt{2}\Gamma(p/2)} \frac{(p-1)}{p} \Sigma^{1/2} \mu + o(p^{-3/2}),
\end{aligned}$$

since $\frac{\mathbf{Z}}{\|\mathbf{Z}\|}$ and $\frac{(\mathbf{Z}^\top \Sigma^{1/2} \mathbf{Z}) \mathbf{Z}}{\|\mathbf{Z}\|^3}$ are symmetric about the origin. We also note that

$$\mathbb{E} \left[\frac{\mathbf{Z} \mathbf{Z}^\top}{\|\mathbf{Z}\|^2} \right] = \frac{1}{p} \mathbf{I}, \text{ and, } \mathbb{E} \left[\frac{\mathbf{Z} (\mathbf{Z}^\top \mathbf{A} \mathbf{Z}) \mathbf{Z}^\top}{\|\mathbf{Z}\|^4} \right] = \alpha(\mathbf{A}, p), \quad \mathbb{E} \left[\frac{\mathbf{Z} (\mathbf{Z}^\top \mathbf{A} \mathbf{Z})^2 \mathbf{Z}^\top}{\|\mathbf{Z}\|^4} \right] = \beta(\mathbf{A}, p)$$

For any general matrix \mathbf{A} , the exact form of $\alpha(\mathbf{A}, p)$ is difficult to obtain. However, one can establish that $\alpha(\mathbf{A}, p)$ lies between $\lambda_p(\mathbf{A})p^{-1}$ and $\lambda_1(\mathbf{A})p^{-1}$, where $\lambda_1 > \lambda_2 > \dots > \lambda_p$ are the eigenvalues of the matrix \mathbf{A} . Similarly, $\lambda_p(\mathbf{A})^2 p^{-1} \leq \beta(\mathbf{A}, p) \leq \lambda_1(\mathbf{A})^2 p^{-1}$.

Now, we use the same Taylor's series expansion on $\mathbf{X} \mathbf{X}^\top / \|\mathbf{X}\|^2$ to obtain

$$\begin{aligned}
\mathbb{E} \left[\frac{\mathbf{X} \mathbf{X}^\top}{\|\mathbf{X}\|^2} \right] &= \mathbb{E} \left[\frac{\mathbf{Z} \mathbf{Z}^\top}{\|\mathbf{Z}\|^2} \right] + 2 \mathbb{E} \left[\left(\frac{\mathbf{I}}{\|\mathbf{Z}\|} - \frac{\mathbf{Z} \mathbf{Z}^\top}{\|\mathbf{Z}\|^3} \right) \Sigma^{1/2} \mu \frac{\mathbf{Z}^\top}{\|\mathbf{Z}\|} \right] + \\
&\quad \mathbb{E} \left[\left(\frac{\mathbf{I}}{\|\mathbf{Z}\|} - \frac{\mathbf{Z} \mathbf{Z}^\top}{\|\mathbf{Z}\|^3} \right) \Sigma^{1/2} \mu \mu^\top \Sigma^{1/2} \left(\frac{\mathbf{I}}{\|\mathbf{Z}\|} - \frac{\mathbf{Z} \mathbf{Z}^\top}{\|\mathbf{Z}\|^3} \right) \right] + 2 \mathbb{E} \left[\left\langle \frac{\mathbf{I}}{\|\mathbf{Z}\|} - \frac{\mathbf{Z} \mathbf{Z}^\top}{\|\mathbf{Z}\|^3}, \Sigma^{1/2} - \mathbf{I} \right\rangle \frac{\mathbf{Z} \mathbf{Z}^\top}{\|\mathbf{Z}\|} \right] + \\
&\quad 2 \mathbb{E} \left[\left\langle \frac{\mathbf{I}}{\|\mathbf{Z}\|} - \frac{\mathbf{Z} \mathbf{Z}^\top}{\|\mathbf{Z}\|^3}, \Sigma^{1/2} - \mathbf{I} \right\rangle \mathbf{Z} \mu^\top \Sigma^{1/2} \left(\frac{\mathbf{I}}{\|\mathbf{Z}\|} - \frac{\mathbf{Z} \mathbf{Z}^\top}{\|\mathbf{Z}\|^3} \right) \right] + \\
&\quad \mathbb{E} \left[\left\langle \frac{\mathbf{I}}{\|\mathbf{Z}\|} - \frac{\mathbf{Z} \mathbf{Z}^\top}{\|\mathbf{Z}\|^3}, \Sigma^{1/2} - \mathbf{I} \right\rangle \mathbf{Z} \mathbf{Z}^\top \left\langle \frac{\mathbf{I}}{\|\mathbf{Z}\|} - \frac{\mathbf{Z} \mathbf{Z}^\top}{\|\mathbf{Z}\|^3}, \Sigma^{1/2} - \mathbf{I} \right\rangle \right] + o(\mathbb{E} \|\mathbf{Z}\|^{-4}) \\
&= \frac{1}{p} \mathbf{I} + \frac{1}{p} \Sigma^{1/2} \mu \mu^\top \Sigma^{1/2} + \frac{\alpha(\Sigma^{1/2} \mu \mu^\top \Sigma^{1/2}, p)}{(p-2)} + 2 \frac{(\text{trace}(\Sigma^{1/2}) - p + 1)}{p} - 2\alpha(\Sigma^{1/2}, p) + \\
&\quad \frac{(\text{trace}(\Sigma^{1/2}) - p + 1)^2}{p} - 2(\text{trace}(\Sigma^{1/2}) - p + 1)\alpha(\Sigma^{1/2}, p) + \beta(\Sigma^{1/2}, p) + o(p^{-2}) \\
&= \frac{1}{p} (\mathbf{I} + \Sigma^{1/2} \mu \mu^\top \Sigma^{1/2}) + \frac{\Sigma^{1/2} \mu \mu^\top \Sigma^{1/2}}{p(p-2)} + \frac{(\text{trace}(\Sigma^{1/2}) - p + 1)(\text{trace}(\Sigma^{1/2}) - p + 3)}{p} \mathbf{I} \\
&\quad - 2 \frac{(\text{trace}(\Sigma^{1/2}) - p + 2)}{p} \Sigma^{1/2} + \frac{1}{p} \Sigma + o(p^{-2})
\end{aligned}$$

□

References

- Mauricio A Alvarez, Lorenzo Rosasco, and Neil D Lawrence. Kernels for vector-valued functions: A review. *arXiv preprint arXiv:1106.6251*, 2011.
- Probal Chaudhuri. On a geometric notion of quantiles for multivariate data. *Journal of the American Statistical Association*, 91(434):862–872, 1996.
- Soudeep Deb and Kaushik Jana. Nonparametric quantile regression for time series with replicated observations and its application to climate data. *arXiv preprint arXiv:2107.02091*, 2021.
- William F Eddy. Ordering of multivariate data. *Computer Science and Statistics: The Interface. L. Billard, Amsterdam: North-Holland*, pages 25–30, 1985.

- John HJ Einmahl and David M Mason. Generalized quantile processes. *The Annals of Statistics*, pages 1062–1078, 1992.
- Alessio Figalli. On the continuity of center-outward distribution and quantile functions. *Nonlinear Analysis*, 177:413–421, 2018.
- Roger Koenker and Gilbert Bassett Jr. Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50, 1978.