Advances in Applied Mathematics 应用数学进展, 2021, 10(11), 3904-3911
Published Online November 2021 in Hans. http://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2021.1011415

Hans汉斯

基于正交投影的部分线性空间自回归模型稳健估计

张　帆，赵培信

重庆工商大学数学与统计学院，重庆

收稿日期：2021年10月19日；录用日期：2021年11月9日；发布日期：2021年11月22日

摘　要

本文考虑部分线性空间自回归模型的稳健估计问题。结合矩阵的正交投影技术和分位数回归方法，对模型的参数提出一种新的基于正交投影的稳健估计方法，并在一些正则条件下研究了估计量的渐进性质。最后通过一些数值模拟研究了所提方法的有限样本性质，模拟结果表明所提出的方法具有较好的稳健性。

关键词

空间自回归模型，部分线性模型，分位数回归，正交投影

Orthogonal Projection Based Robust Estimation of Partial Linear Spatial Autoregressive Model

Fan Zhang, Peixin Zhao

College of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing

Received: Oct. 19th, 2021; accepted: Nov. 9th, 2021; published: Nov. 22nd, 2021

Abstract

In this paper, we consider the robust estimation of partial linear spatial autoregressive models. Combining the orthogonal projection technique of matrix and quantile regression method, a new robust estimation method based on orthogonal projection is proposed for the parameters of the model, and the asymptotic properties of the estimator are studied under some regular conditions. Finally, the finite sample properties of the proposed method are studied through some numerical simulations. The simulation results show that the proposed method has good robustness.

文章引用：张帆，赵培信．基于正交投影的部分线性空间自回归模型稳健估计[J]．应用数学进展，2021，10(11): 3904-3911. DOI: 10.12677/aam.2021.1011415

Keywords

Spatial Autoregressive Model, Partially Linear Model, Quantile Regression, Orthogonal Projection

1. 引言

空间自回归模型广泛应用在经济学、政治学和公共健康学等众多的领域，其用于解决地理单元上的空间交互效应问题。关于空间自回归模型的一些早期研究可参见文献 Anselin (1988) [1]，Anselin 和 Bera (1998) [2]，LeSage (1999) [3]以及 LeSage 和 Pace (2009) [4]等。近几年来，有大量的文章将关注点聚焦在空间计量经济模型的统计推断研究上。例如 Lee (2007) [5]研究将广义矩估计方法应用到空间自回归模型里，Lee (2004) [6]研究了空间自回归模型的拟极大似然估计量的渐近性。Lee 和 Yu (2010) [7]提出了带有空间滞后和空间扰动的空间自回归面板模型的极大似然估计。Xu 和 Lee (2015) [8]考虑了工具变量和极大似然估计量在空间自回归模型中的应用。Qu 和 Lee (2015) [9]提出了三条对于含有内生空间权重矩阵的空间自回归模型的估计方法，包括两阶段工具变量法，拟极大似然估计法，以及广义矩估计法。

然而在实际应用中，线性结构的空间自回归模型并不足以刻画潜在的变量结构方式。另外，仅仅依靠非参数回归方法又有可能会面临“维数灾祸”的问题。为此，最近一些降维的半参数空间自回归模型越来越受到关注，其中部分线性空间自回归模型作为半参数空间自回归模型的主要模型之一已被部分学者研究，并提出了一些统计推断方法。例如，Su 和 Jin (2010) [10]针对部分线性空间自回归模型提出了拟极大似然估计的方法，Cheng 和 Chen (2019) [11]研究了针对部分线性单指标空间自回归模型的广义矩估计，Du 等人(2018) [12]对部分线性可加空间自回归模型进行了研究。

但是以往的文献主要是基于均值回归框架下研究部分线性空间自回归模型的统计推断问题，而均值回归往往会受到数据异常值的影响。基于此，本文主要研究部分线性空间自回归模型的稳健估计问题。具体地，基于分位数回归方法以及矩阵的 QR 分解技术，本文对部分线性空间自回归模型提出了一种基于正交投影的稳健估计方法。另外在一些正则条件下，研究了所得估计的渐近性质，并且通过数据模拟验证了所提出的估计方法具有较好的稳健性。

2. 方法论及主要结果

设我们具有一个容量为 n 的样本，并记 y_i 是因变量，X_i 是 $p\times 1$ 维自变量，w_{ij} 是空间权重矩阵 W 的第 (i,j) 元素。参数 ρ 是空间滞后因变量 Wy 的一个效应系数，β 是 $p\times 1$ 维的参数向量，u_i 是一维的光滑变量。那么部分线性空间自回归模型为

$$y_i=\rho\sum_{j=1}^{n}w_{ij}y_j+X_i^{\mathrm{T}}\beta+g(u_i)+\varepsilon_i=\rho d_i+X_i^{\mathrm{T}}\beta+g(u_i)+\varepsilon_i \tag{1}$$

其中 $d_i=\sum_{j=1}^{n}w_{ij}y_j$，$X_i=(X_{i1},\cdots,X_{ip})^{\mathrm{T}}$。

由于内生变量 d_i 的存在，我们采用工具变量分位数回归(IVQR)的方法来消除偏差。内生变量 d_i 和一组工具变量 ω_i 相关，而 ω_i 与 ε_i 无关。这样我们就可以定义条件工具分位数为：

$$Q_\tau\left(y_i \mid \mathcal{F}_{-i}, X_i, u_i\right)=\rho(\tau) d_i+X_i^{\mathrm{T}} \beta(\tau)+g\left(u_i, \tau\right)+\omega_i \gamma(\tau) \tag{2}$$

其中 $Q_\tau\left(y_i \mid \mathcal{F}_{-i}, X_i, u_i\right)$ 是给定 $\mathcal{F}_{-i}, X_i, u_i$ 值下的 y_i 的 τ 分位数，$\mathcal{F}_{-i}$ 是 σ 域 $\left\{y_j: j \neq i\right\}$，$\gamma$ 是工具变量 ω_i 对应的系数，$Q_\tau\left(\varepsilon_i \mid \mathcal{F}_{-i}, X_i, u_i\right)=0$。

我们采用 B 样条来进行估计。为了不失一般性，我们假设 $U_i \in[0,1]$ 对所有的 i 都成立。我们采用阶数为 $h+1$ 的 B 样条基函数来估计 $g\left(u_i, \tau\right)$，考虑一列正整数序列 $\left\{k_n\right\}$，$n \geq 1$，将 $[0,1]$ 分成 k_n 个近似均匀的区间。令 $\pi(u)=\left(B_1(u), \cdots, B_L(u)\right)^{\mathrm{T}}$ 表示一组 B 样条基函数，其中 $L=k_n+h+1$，则 $g\left(u_i, \tau\right)$ 便可被这组规范化的 B 样条基函数线性表示为

$$g(u, \tau) \approx \sum_{s=1}^L B_s(u) \theta_s(\tau)=\pi(u)^{\mathrm{T}} \theta(\tau)$$

此处，$\theta(\tau)=\left(\theta_1(\tau), \cdots, \theta_L(\tau)\right)^{\mathrm{T}}$ 是样条系数向量。关于 B 样条基函数构建的细节，读者可参考 Schumaker (1981) [13]。在 B 样条展开的基础上，模型(2)可近似为

$$\begin{aligned} Q_\tau\left(y_i \mid \mathcal{F}_{-i}, X_i, Z_i, u_i\right) & \approx \rho(\tau) d_i+\sum_{j=1}^p X_{i j} \beta_j(\tau)+\sum_{s=1}^L B_s\left(u_i\right) \theta_s(\tau)+\omega_i \gamma(\tau) \\ & =\rho(\tau) d_i+X_i^{\mathrm{T}} \beta(\tau)+\pi\left(u_i\right)^{\mathrm{T}} \theta(\tau)+\omega_i \gamma(\tau) \end{aligned} \tag{3}$$

模型(3)的矩阵形式为

$$y=\rho(\tau) W y+X \beta(\tau)+\Pi \theta(\tau)+\Omega \gamma(\tau)+\varepsilon \tag{4}$$

其中，$X=\left(X_1, \cdots, X_n\right)^{\mathrm{T}}$，$\Pi=\left(\pi\left(u_1\right), \cdots, \pi\left(u_n\right)\right)^{\mathrm{T}}$，$\Omega=\left(\omega_1, \cdots, \omega_n\right)^{\mathrm{T}}$，$y=\left(y_1, \cdots, y_n\right)$，$\varepsilon=\left(\varepsilon_1, \cdots, \varepsilon_n\right)^{\mathrm{T}}$ 对 Π 进行 QR 分解得

$$\Pi=Q R=\left(Q_1, Q_2\right)\left(R_1^{\mathrm{T}}, 0\right)^{\mathrm{T}}=Q_1 R_1,$$

Q_1 是 $n \times L$ 阶矩阵，R_1 是 $L \times L$ 阶方阵，Q_2 是 $n \times(n-L)$ 阶矩阵，$Q_2^{\mathrm{T}} \Pi=Q_2 Q_1 R_1=0$。在(4)式两边同时左乘 Q_2^{T}，得

$$Q_2^{\mathrm{T}} y=\rho(\tau) Q_2^{\mathrm{T}} W y+Q_2^{\mathrm{T}} X \beta(\tau)+Q_2^{\mathrm{T}} \Omega \gamma(\tau)+Q_2^{\mathrm{T}} \varepsilon \tag{5}$$

其中 $Q_2^{\mathrm{T}} y=\left(\tilde{y}_1, \cdots, \tilde{y}_{n-L}\right)$，$Q_2^{\mathrm{T}} W y=\left(\tilde{d}_1, \cdots, \tilde{d}_{n-L}\right)^{\mathrm{T}}$，$Q_2^{\mathrm{T}} \Omega=\left(\tilde{\omega}_1, \cdots, \tilde{\omega}_{n-L}\right)^{\mathrm{T}}$，$Q_2^{\mathrm{T}} X=\left(\tilde{X}_1, \cdots, \tilde{X}_{n-L}\right)^{\mathrm{T}}$，$Q_2^{\mathrm{T}} \varepsilon=\left(\tilde{\varepsilon}_1, \cdots, \tilde{\varepsilon}_{n-L}\right)^{\mathrm{T}}$，于是

$$\tilde{y}_i=\rho(\tau) \tilde{d}_i+\tilde{X}_i^{\mathrm{T}} \beta(\tau)+\tilde{\omega}_i \gamma(\tau)+\tilde{\varepsilon}_i \tag{6}$$

接下来我们定义目标函数：

$$R_{IV}(\tau, \rho, \beta, \gamma)=\sum_{i=1}^{n-L} \rho_\tau\left(\tilde{y}_i-\rho \tilde{d}_i-\tilde{X}_i^{\mathrm{T}} \beta-\tilde{\omega}_i \gamma\right) \tag{7}$$

其中 $\rho_\tau(u):=(\tau-I(u<0)) u$，根据 Chernozhukov 和 Hansen (2006, 2008) [14] [15]以及 Galvao (2011) [16]，假设工具变量 ω_i 可以获得，则我们可以依照以下三步推导出 IVQR 估计量。

第一步，给定一个分位数 τ，定义一组合适的 $\left\{\rho_j, j=1, \cdots, J ;\left|\rho_j\right|<1\right\}$，然后最小化目标函数 R_{IV}，获得分位数估计量 β, γ。

$$(\hat{\beta}(\rho, \tau), \hat{\gamma}(\rho, \tau))=\underset{\beta, \gamma}{\arg \min } R_{IV}(\tau, \rho, \beta, \gamma) \tag{8}$$

第二步，在 $\left\{\rho_j, j=1, \cdots, J ;\left|\rho_j\right|<1\right\}$ 中选出一个 $\hat{\rho}_\tau$，使得定义在 ζ

$$\rho_\tau = \arg\min_{\rho \in \mathcal{R}} \left\{ \hat{\gamma}_\tau^{\mathrm{T}} A \hat{\gamma}_\tau \right\} \tag{9}$$

的加权距离函数最接近零：其中 A 是一个正定矩阵，$\mathcal{R} = [-1,1]$。

第三步，β 的估计可以得到，为 $\hat{\beta}(\hat{\rho}_\tau, \tau)$。

在本篇论文中，我们运用三阶 B 样条进行估计。针对目标函数(7)式，我们根据 Schwarz-type 信息准则来选取 k_n 个节点：

$$SIC(k_n) = \log \sum_{i=1}^{n-L} \rho_\tau \left(\tilde{y}_i - \hat{\rho}_{(k_n)} \tilde{d}_i - \tilde{X}_i^{\mathrm{T}} \hat{\beta}_{(k_n)} - \tilde{\omega}_i \hat{\gamma}_{(k_n)} \right) + \frac{\log n}{2n} (2 + p + L)$$

其中 $\hat{\rho}_{(k_n)}$，$\hat{\beta}_{(k_n)}$，$\hat{\gamma}_{(k_n)}$ 分别为有 k_n 个节点的第 τ 分位点上的估计量，更多细节可见 Kim (2003) [17]。另外，为了得到 IVQR 估计量，对于内生变量 $D = Wy$ 我们需要一组工具变量。在本文中，采用 WX 作为工具变量矩阵。

为了使用 IVQR 估计量，下面需要引入几条正则性条件。

正则条件 1

1) (y_i, u_i, D_i) 是独立同分布的，y_i 的条件分布函数为 F，$i = 1, \cdots, n$。

2) 对每一个 i，ε_i 的密度函数 f_i 是有界的，该函数在零点处可微，且一阶导数有界。

正则条件 2

1) 对于所有的 $\tau \in \mathcal{T}$，$(\rho(\tau), \beta(\tau))$ 在集合 $\mathcal{R} \times \mathcal{B}$ 中，且 $\mathcal{R} \times \mathcal{B}$ 既是紧的又是凸的。

2) 令

$$\Phi(\rho, \beta, \gamma, \tau) = E\left[\left(\tau - I\left(\tilde{y} < \tilde{D}\rho + \tilde{X}\beta + \tilde{\Omega}\gamma \right) \right) \bar{X} \right], \tag{10}$$

$$\Phi(\rho, \beta, \tau) = E\left[\left(\tau - I\left(\tilde{y} < \tilde{D}\rho + \tilde{X}\beta \right) \right) \bar{X} \right], \tag{11}$$

其中 $\bar{X} = (\tilde{X}, \tilde{\Omega})$，$\tilde{D} = Q_2^{\mathrm{T}} W y$，$\tilde{\Omega} = (\tilde{\omega}_1, \cdots, \tilde{\omega}_{n-L})^{\mathrm{T}}$。雅可比矩阵 $\dfrac{\partial \Phi(\rho, \beta, \tau)}{\partial(\rho, \beta)}$ 和 $\dfrac{\partial \Phi(\rho, \beta, \gamma, \tau)}{\partial(\beta, \gamma)}$ 是连续的且为满秩。参数空间 $\mathcal{R} \times \mathcal{B}$ 是连通集且在此集合的像在映射 $(\rho, \beta) \mapsto \Phi(\rho, \beta, \tau)$ 下是连通的。

3) 定义 $\Omega = \mathrm{diag}\{ f_i(\xi_i(\tau)) \}$，其中 $\xi_i(\tau) = \rho(\tau)\tilde{d}_i + \tilde{X}_i^{\mathrm{T}} \beta(\tau) + \tilde{\omega}_i \gamma(\tau)$。令 $\eta = (\beta^{\mathrm{T}}, \gamma)^{\mathrm{T}}$。则下列矩阵均为正定矩阵：

$$J_\eta = \lim_{n \to \infty} \frac{1}{n} \bar{X}^{\mathrm{T}} \Omega \bar{X}, \tag{12}$$

$$J_\rho = \lim_{n \to \infty} \frac{1}{n} \bar{X}^{\mathrm{T}} \Omega \tilde{D}, \tag{13}$$

$$S = \lim_{n \to \infty} \frac{\tau(1-\tau)}{n} \bar{X}^{\mathrm{T}} \bar{X} \tag{14}$$

其中 $\left[J_\beta^{\mathrm{T}}, J_\gamma^{\mathrm{T}} \right]$ 是 J_η^{-1} 的一个合适划分。因此，J_η 是可逆的且 $J_\rho^{\mathrm{T}} H J_\rho$ 也是可逆的。

4) $\max \|y_i\| = O(\sqrt{n})$，$\max \|X_i\| = O(\sqrt{n})$，$\max \|\omega_i\| = O(\sqrt{n})$。

定理 1：在正则条件 1 和 2 下，对任意给定的 $\tau \in (0,1)$，$\hat{\vartheta}(\tau) = \left(\hat{\rho}(\tau), \hat{\beta}^{\mathrm{T}}(\tau) \right)^{\mathrm{T}}$ 依分布收敛于高斯分布，即

$$\sqrt{n} \left(\hat{\vartheta}(\tau) - \vartheta(\tau) \right) \xrightarrow{d} N\left(0, J^{\mathrm{T}} S J \right) \tag{15}$$

其中 $J = (K^{\mathrm{T}}, L_1^{\mathrm{T}})$，$L_1 = J_\beta^{\mathrm{T}} M$，$M = I - J_\rho K$，$K = \left(J_\rho^{\mathrm{T}} H J_\rho \right)^{-1} J_\rho^{\mathrm{T}} H$，对任意给定的 $\tau \in (0,1)$，结合定理

1 可知 $\beta(\tau)$ 的 $100(1-\alpha)\%$ 置信区间可表为

$$\left[\hat{\beta}(\tau)-\frac{Z_{\alpha/2}}{n}\sigma_\beta,\hat{\beta}(\tau)+\frac{Z_{\alpha/2}}{n}\sigma_\beta\right]$$

其中 $\sigma_\beta=\left(\Lambda_{\beta 11}^{1/2},\cdots,\Lambda_{\beta pp}^{1/2}\right)^\top$， $\Lambda_{\beta ii}$ 是对角矩阵 Λ_β 的第 i 个对角元素， $\Lambda_\beta=L_1SL_1$ 。

3. 数值模拟研究

在本节中，我们利用蒙特卡洛方法来研究本文提出的估计方法的有限样本性质。在模拟过程中，样本容量分别取 $n=200,400,600$ 三种情况，并且在每个样本下，实验重复 1000 次。对估计量进行的分位数回归时分位点分别取为 $\tau=0.25,0.5,0.75$ 。数据从如下模型产生：

$$y_i=\rho\sum_{j=1}^{n}w_{ij}y_j+X_i\beta+g(u_i)+\varepsilon_i,\quad i=1,\cdots,n, \tag{16}$$

其中 $\rho=0.5$， $\beta=2$， $g(u)=\sin(2\pi u)$， $\varepsilon_i=e_i-F^{-1}(\tau)$，$F$ 是 e_i 的分布函数，因此随机误差 ε_i 是以 τ 分位点为中心。为了说明本文提出方法的稳健性，模型误差 e_i 分别取为对称分布 $N(0,0.5)$ 和非对称分布 $\chi^2(1)$ 两种情况。这里协变量 u_i,X_i 分别服从 $U(0,2)$， $N(0,1)$ 分布，另外类似 Dai 等人(2016) [18]，空间权重矩阵 $W=(w_{ij})$ 的生成机制为 $w_{ij}=r^{|i-j|}I(i\neq j)$，其中 $r=0.3$， $1\le i$， $j\le n$ 。

在本文的模拟研究中，我们把本文提出的基于分位数回归得到的估计结果与 Wang 等人(2016) [19]提出的均值回归得到的估计结果进行比较。表 1 给出了基于这两种方法，在不同情况下估计量的偏差和标准差基于 1000 次重复实验的平均值。由表 1 可以得到以下结论：

1) 对任一给定的分位点 τ，基于本文方法给出的估计偏差和标准差均随着样本量的增加而减小。并且对任意给定的样本量n，本文的方法在两种误差分布下给出模拟结果是非常类似的，这也表明本文提出的估计方法具有较好的稳健性。

2) 对任一给定的分位点 τ，当误差分布为对称分布时，本文方法给出的模拟结果与基于文献Ahmad 等人(2005) [20]的均值回归方法给出的结果是类似的。但是当误差分布为非对称分布，即数据中含有部分异常值时，本文方法给出的模拟结果明显优于均值回归主法给出的模拟结果。

Table 1. Simulation results of parameter components based on different estimation methods
表 1. 基于不同估计方法对参数分量的模拟结果

Err. Dist.	n	分位数回归						均值回归	
		Bias	SD	Bias	SD	Bias	SD	Bias	SD
		$\tau=0.25$	$\tau=0.25$	$\tau=0.5$	$\tau=0.5$	$\tau=0.75$	$\tau=0.75$	$\tau=0.5$	$\tau=0.5$
$N(0,0.5)$	200	0.0057	0.235	0.0054	0.205	0.0052	0.207	0.0051	0.203
	400	−0.0011	0.134	0.0015	0.138	−0.0014	0.136	0.0014	0.135
	600	−0.0007	0.112	−0.0008	0.110	−0.0008	0.118	−0.0009	0.119
$\chi^2(1)$	200	0.0052	0.209	0.0053	0.204	0.0052	0.202	0.0393	0.567
	400	−0.0017	0.134	−0.0013	0.131	−0.0014	0.135	−0.0335	0.528
	600	0.0005	0.122	−0.0008	0.112	0.0007	0.110	−0.0292	0.502

4. 证明

为了证明定理 1，我们要先引入一个引理。

引理 1：定义 $\tilde{\varepsilon}_i(\tau)=\tilde{y}_i-\xi_i(\tau)$，令 $\vartheta^*=(\rho,\beta,\gamma)$ 是一组定义在 $\mathcal{R}\times\mathcal{B}\times\mathcal{Z}$ 空间的参数向量，并记

$$\delta=\begin{pmatrix}\delta_\rho\\ \delta_\beta\\ \delta_\gamma\end{pmatrix}=\begin{pmatrix}\sqrt{n}\left(\hat{\rho}(\tau)-\rho(\tau)\right)\\ \sqrt{n}\left(\hat{\beta}(\tau)-\beta(\tau)\right)\\ \sqrt{n}\left(\hat{\gamma}(\tau)-\gamma(\tau)\right)\end{pmatrix}\tag{17}$$

则在正则条件 1 和 2 下，有

$$\begin{aligned}\sup_{\vartheta^*}\frac{1}{n}\Bigg|\sum_{i=1}^{n-L}\Bigg[&\rho_\tau\left(\tilde{\varepsilon}_i(\tau)-\frac{\tilde{d}_i\delta_\rho}{\sqrt{n}}-\frac{\tilde{X}_i\delta_\beta}{\sqrt{n}}-\frac{\tilde{\omega}_i\delta_\gamma}{\sqrt{n}}\right)-\rho_\tau\left(\tilde{\varepsilon}_i(\tau)\right)\\ &-E\left[\rho_\tau\left(\tilde{\varepsilon}_i(\tau)-\frac{\tilde{d}_i\delta_\rho}{\sqrt{n}}-\frac{\tilde{X}_i\delta_\beta}{\sqrt{n}}-\frac{\tilde{\omega}_i\delta_\gamma}{\sqrt{n}}\right)-\rho_\tau\left(\tilde{\varepsilon}_i(\tau)\right)\right]\Bigg]\Bigg|\end{aligned}\tag{18}$$

具体证明详见 Dai *et al*. (2016) [18]。

定理 1 证明

接下来，根据 Chernozhukov 和 Hansen (2006) [14]的结论，$\vartheta^*(\tau)=\left(\rho(\tau),\beta(\tau)\right)$ 对于每一个 τ 都有唯一解。

为了证明参数的连续性，在正则条件 1 和 2 下，我们有 $\hat{\vartheta}^*=\vartheta^*+o_p(1)$。令

$$\mathcal{P}:\vartheta^*\mapsto\rho_\tau\left(\tilde{y}-\rho\tilde{D}-\tilde{X}\beta\right)$$

其中 $\mathcal{P}$ 是连续的。在引理 1 的条件下，对于 $\vartheta^*=(\rho,\beta,\gamma)$ 我们有 $\left\|\hat{\vartheta}^*(\rho,\tau)-\vartheta^*(\rho,\tau)\right\|\xrightarrow{P}0$。根据 Dai (2016) [18]的结论，我们有 $\left\|\hat{\rho}(\tau)-\rho(\tau)\right\|\xrightarrow{P}0$。因此我们可以得到 $\left\|\hat{\beta}\left(\hat{\rho}(\tau),\tau\right)-\beta(\tau)\right\|\xrightarrow{P}0$，同时也得出 $\left\|\hat{\gamma}\left(\hat{\rho}(\tau),\tau\right)-0\right\|\xrightarrow{P}0$，所以，$\left\|\hat{\vartheta}^*(\tau)-\vartheta^*(\tau)\right\|\xrightarrow{P}0$。

当 $\hat{\rho}(\tau)\xrightarrow{P}\rho(\tau)$ 时，我们可以将(8)式的目标函数写成

$$R_{IV}=\sum_{i=1}^{n-L}\left[\rho_\tau\left(\tilde{\varepsilon}_i(\tau)-\frac{\tilde{d}_i\delta_\rho}{\sqrt{n}}-\frac{\tilde{X}_i\delta_\beta}{\sqrt{n}}-\frac{\tilde{\omega}_i\delta_\gamma}{\sqrt{n}}\right)-\rho_\tau\left(\tilde{\varepsilon}_i(\tau)\right)\right]$$

其中 $\tilde{\varepsilon}_i(\tau)=\tilde{y}_i-\tilde{\xi}_i(\tau)$，$\tilde{\xi}_i(\tau)=\rho(\tau)\tilde{d}_i+\tilde{X}_i^{\mathrm{T}}\beta(\tau)+\tilde{\omega}_i\gamma(\tau)$。令 $\varphi_\tau(u)=\tau-I(u<0)$，

$$G\left(\delta_\rho,\delta_\beta,\delta_\gamma\right)=\frac{-1}{\sqrt{n}}\sum_{i=1}^{n-L}\varphi_\tau\left(\tilde{\varepsilon}_i(\tau)-\frac{\tilde{d}_i\delta_\rho}{\sqrt{n}}-\frac{\tilde{X}_i\delta_\beta}{\sqrt{n}}-\frac{\tilde{\omega}_i\delta_\gamma}{\sqrt{n}}\right)$$

于是 $\sup\left\|G\left(\delta_\rho,\delta_\beta,\delta_\gamma\right)-G(0,0,0)-E\left[G\left(\delta_\rho,\delta_\beta,\delta_\gamma\right)-G(0,0,0)\right]\right\|=o_p(1)$。将 G 展开得

$$E\left[G\left(\delta_\rho,\delta_\beta,\delta_\gamma\right)-G(0,0,0)\right]=\frac{1}{\sqrt{n}}\sum_{i=1}^{n-L}\bar{X}_i^{\mathrm{T}}f_i\left(\xi_i(\tau)\right)\left[\frac{\tilde{d}_i\delta_\rho}{\sqrt{n}}+\frac{\tilde{X}_i\delta_\beta}{\sqrt{n}}+\frac{\tilde{\omega}_i\delta_\gamma}{\sqrt{n}}\right]+o_p(1)\tag{19}$$

其中 $\bar{X}=\left(\tilde{X},E\right)$。显然 $G\left(\hat{\delta}_\rho,\hat{\delta}_\beta,\hat{\delta}_\gamma\right)\to0$，$E\left[G\left(\delta_\rho,\delta_\beta,\delta_\gamma\right)-G(0,0,0)\right]=-G(0,0,0)$，则最后的等式可表示为

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n-L}\bar{X}_i^{\mathrm{T}}\varphi_\tau\left(\varepsilon_i(\tau)\right)=\frac{1}{\sqrt{n}}\sum_{i=1}^{n-L}\bar{X}_i^{\mathrm{T}}f_i\left(\xi_i(\tau)\right)\left[\frac{\tilde{d}_i\delta_\rho}{\sqrt{n}}+\frac{\tilde{X}_i\delta_\beta}{\sqrt{n}}+\frac{\tilde{\omega}_i\delta_\gamma}{\sqrt{n}}\right]+o_p(1)\tag{20}$$

令 $\delta_\eta = \left(\delta_\beta^{\mathrm{T}}, \delta_\gamma\right)^{\mathrm{T}}$，上式可写为

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n-L}\bar{X}_i^{\mathrm{T}} f_i\left(\xi_i(\tau)\right)\left[\frac{\tilde{d}_i\delta_\rho}{\sqrt{n}}+\frac{\bar{X}_i\delta_\eta}{\sqrt{n}}\right]=\frac{1}{\sqrt{n}}\sum_{i=1}^{n-L}\bar{X}_i^{\mathrm{T}}\varphi_\tau\left(\varepsilon_i(\tau)\right)$$

运用更简单的记号，上式可表示为

$$J_\rho\delta_\rho + J_\eta\delta_\eta = J_\phi$$

其中 $J_\eta = \lim\limits_{n\to\infty}\frac{1}{n}\bar{X}^{\mathrm{T}}\Omega\bar{X}$， $J_\rho = \lim\limits_{n\to\infty}\frac{1}{n}\bar{X}^{\mathrm{T}}\Omega\tilde{D}$， J_ϕ 是一个零均值的随机向量，其协方差为 $\tau(1-\tau)\bar{X}^{\mathrm{T}}\bar{X}$，$\Omega = \mathrm{diag}\left\{f_i\left(\xi_i(\tau)\right)\right\}$。

令 $\left[J_\beta^{\mathrm{T}}, J_\zeta^{\mathrm{T}}\right]$ 是 J_η^{-1} 的一个合适划分(如 Galvao (2011)和 Chernozhukov 以及 Hansen (2006) [14] [16])使得 $\hat{\delta}_\beta = J_\beta^{\mathrm{T}}\left(J_\phi - J_\rho\delta_\rho\right)$， $\hat{\delta}_\eta = J_\eta^{\mathrm{T}}\left(J_\phi - J_\rho\delta_\rho\right)$。令 $H = J_\gamma^{\mathrm{T}} A J_\gamma$， 则 $\hat{\delta}_\beta = KJ_\phi$， 其中 $K = \left(J_\rho^{\mathrm{T}} H J_\rho\right)^{-1} J_\rho^{\mathrm{T}} H$。将它带入之前的表达式中，则 $\hat{\delta}_\eta = J_\eta^{\mathrm{T}}\left(J_\phi - J_\rho\delta_\rho\right) = J_\eta^{\mathrm{T}}\left(I - J_\rho\left(J_\rho^{\mathrm{T}} H J_\rho\right)^{-1} J_\rho^{\mathrm{T}} H\right) J_\phi = J_\eta^{\mathrm{T}} M J_\phi$， 其中 $M = I - J_\rho\left(J_\rho^{\mathrm{T}} H J_\rho\right)^{-1} J_\rho^{\mathrm{T}} H$。由于 $J_\rho J_\eta$ 可逆，所以 $\hat{\delta}_\eta = \mathbf{0}\times O_p(1) + o_p(1)$。同理可得 $\hat{\delta}_\beta = J_\beta^{\mathrm{T}} M J_\phi$。在正规条件下，我们有

$$\begin{pmatrix}\hat{\delta}_\rho\left(\rho_n,\tau\right)\\ \hat{\delta}_\beta\left(\rho_n,\tau\right)\end{pmatrix}=\begin{pmatrix}\sqrt{n}\left(\hat{\rho}\left(\rho_n,\tau\right)-\rho(\tau)\right)\\ \sqrt{n}\left(\hat{\beta}\left(\rho_n,\tau\right)-\beta(\tau)\right)\end{pmatrix}\rightsquigarrow \mathcal{N}\left(\mathbf{0}, J^{\mathrm{T}} S J\right)$$

其中 $J = \left(K^{\mathrm{T}}, L_1^{\mathrm{T}}\right)$， $L_1 = J_\beta^{\mathrm{T}} M$。

参考文献

[1] Anselin, L. (1988) Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, Dordrecht. https://doi.org/10.1007/978-94-015-7799-1

[2] Anselin, L. and Bera, A.K. (1998) Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In: Ullah, A., Ed., *Handbook of Applied Economic Statistics*, CRC Press, New York, 237-290.

[3] LeSage, J.P. (1999) The Theory and Practice of Spatial Econometrics. University of Toledo, Toledo.

[4] LeSage, J.P. and Pace, R.K. (2009) Introduction to Spatial Econometrics. Chapman & Hall/CRC, New York. https://doi.org/10.1201/9781420064254

[5] Lee, L.F. (2007) GMM and 2SLS Estimation of Mixed Regressive, Spatial Autoregressive Models. *Journal of Econometrics*, **137**, 489-514. https://doi.org/10.1016/j.jeconom.2005.10.004

[6] Lee, L.F. (2004) Asymptotic Distributions of Quasi-Maximum Likelihood Estimators for Spatial Autoregressive Models. *Econometrica*, **72**, 1899-1925. https://doi.org/10.1111/j.1468-0262.2004.00558.x

[7] Lee, L.F. and Yu, J.H. (2010) Estimation of Spatial Autoregressive Panel Data Models with Fixed Effects. *Journal of Econometrics*, **154**, 165-185. https://doi.org/10.1016/j.jeconom.2009.08.001

[8] Xu, X.B. and Lee, L.F. (2015) A Spatial Autoregressive Model with a Nonlinear Transformation of the Dependent Variable. *Journal of Econometrics*, **186**, 1-18. https://doi.org/10.1016/j.jeconom.2014.12.005

[9] Qu, X. and Lee, L.F. (2015) A Spatial Autoregressive Model with a Nonlinear Transformation of the Dependent Variable. *Journal of Econometrics*, **184**, 209-232. https://doi.org/10.1016/j.jeconom.2014.08.008

[10] Su, L.J. and Jin, S.N. (2010). Profile Quasi-Maximum Likelihood Estimation of Partially Linear Spatial Autoregressive Models. *Journal of Econometrics*, **157**, 18-33. https://doi.org/10.1016/j.jeconom.2009.10.033

[11] Cheng, S.L., Chen, J.B. and Liu, X. (2019) GMM Estimation of Partially Linear Single-Index Spatial Autoregressive Model. *Spatial Statistics*, **31**, Article ID: 100354. https://doi.org/10.1016/j.spasta.2019.04.002

[12] Du, J., Sun, X.Q., Cao, R.Y. and Zhang, Z.Z. (2018) Statistical Inference for Partially Additive Spatial Autoregressive Models. *Spatial Statistics*, **25**, 52-67. https://doi.org/10.1016/j.spasta.2018.04.008

[13] Chernozhukov, V. and Hansen, C. (2008) Instrumental Variable Quantile Regression: A Robust Inference Approach. *Journal of Econometrics*, **142**, 379-398. https://doi.org/10.1016/j.jeconom.2007.06.005

[14] Chernozhukov, V. and Hansen, C. (2006) Instrumental Quantile Regression Inference for Structural and Treatment Effect Models. *Journal of Econometrics*, **132**, 491-525. https://doi.org/10.1016/j.jeconom.2005.02.009

[15] Chernozhukov, V. and Hansen, C. (2008) Instrumental Variable Quantile Regression: A Robust Inference Approach. *Journal of Econometrics*, **142**, 379-398. https://doi.org/10.1016/j.jeconom.2007.06.005

[16] Galvao, A.F. (2011) Quantile Regression for Dynamic Panel Data with Fixed Effects. *Journal of Econometrics*, **164**, 142-157. https://doi.org/10.1016/j.jeconom.2011.02.016

[17] Kim, M.O. (2003) Quantile Regression in a Varying Coefficient Model. Ph.D. Dissertation, University of Illinois Urbana-Champaign, Champaign.

[18] Dai, X.W., Jin, L.B., Shi, A.Q. and Shi, L. (2016) Outlier Detection and Accommodations in General Spatial Models. *Statistical Methods and Applications*, **25**, 453-475. https://doi.org/10.1007/s10260-015-0348-1

[19] Dai, X.W., Li, S.Y. and Tian, M.Z. (2016) Quantile Regression for Partially Linear Varying Coefficient Spatial Autoregressive Models.

[20] Wang, H.X., Zhu, Z.Y. and Zhou, J.H. (2009) Quantile Regression in Partially Linear Varying Coefficient Models. *Annals of Statistics*, **37**, 3841-3866. https://doi.org/10.1214/09-AOS695