第27卷第4期
2021年11月
广西民族大学学报(自然科学版)
JOURNAL OF GUANGXI UNIVERSITY FOR NATIONALITIES
(Natural Science Edition)
Vol.27 No.4
November 2021

含空间自回归误差的空间自回归模型的经验欧氏似然推断*

唐　洁
(广西师范大学 数学与统计学院,广西 桂林 541004)

数学与计算机技术

摘　要: 研究了含空间自回归误差的空间自回归模型(SARAR模型)的经验欧氏似然推断问题,利用经验欧氏似然方法,通过鞅差序列处理SARAR模型的估计方程中出现的二次型形式,构造出SARAR模型的经验欧氏似然比统计量,还证明了该经验欧氏似然比统计量是渐近卡方分布。

关键词: SARAR模型;经验欧氏似然;鞅差序列
中图分类号: O212.1　**文献标识码:** A　**文章编号:** 1673-8462(2021)04-0070-05
DOI:10.16177/j.cnki.gxmzzk.2021.04.002

0　引言

含空间自回归误差的空间自回归模型(Spatial Autoregressive Model with Spatial Autoregressive Disturbances,简记为SARAR):

$$Y_n=\rho_1 W_n Y_n+X_n\beta+u_{(n)},\ u_{(n)}=\rho_2 M_n u_n+\epsilon_{(n)} \tag{1}$$

其中,n是空间样本数量;$\rho_j,j=1,2$是空间自回归系数且$\left|\rho_j\right|<1,j=1,2$;$\beta_{k\times 1}$是相应回归系数向量;$X_n=\left(x_1,x_2,\cdots,x_n\right)'$是$n\times k$数据矩阵,包括$k$列解释变量;$Y_n=\left(y_1,y_2,\cdots,y_n\right)'$是$n\times 1$维被解释变量;$W_n$和$M_n$是已知的$n\times n$空间权重矩阵(非随机);$\epsilon_{(n)}$是空间误差模型$n\times 1$的误差向量,$E\epsilon_{(n)}=0$,$\mathrm{Var}(\epsilon_{(n)})=\sigma^2 I_n$。

SARAR模型是更一般的空间计量模型,它对存在空间依赖性的数据有较好的解释作用,无论是滞后项存在依赖性还是扰动项存在空间依赖性。它将存在空间误差效应的空间误差模型(SEM)与空间滞后效应的空间自回归模型(SAR)结合起来,分别对应于$\rho_1=0$与$\rho_2=0$的情形。

经验似然方法是由Owen在文献[1]中提出的一种参数推断方法。通过经验似然方法研究SARAR模型则是文献[2]的主要工作。但是经验似然方法在求解过程中会出现没有显示解的情况,且计算复杂。为了解决这些问题,Owen在文献[3]中提出用经验欧氏似然来代替经验似然。而罗旭在文献[4]中,就系统地研究了经验欧氏似然,发现了可以很好地解决经验似然中的棘手问题,并且经验欧氏似然也同样拥有大样本性质。基于此,本文通过经验欧氏似然方法来研究SARAR模型。

* 收稿日期:2021-07-27.
基金项目:广西研究生教育创新计划项目(YJSCXP202104)。
作者简介:唐洁(1995-),女,湖南人,广西师范大学数学与统计学院2020级统计学专业硕士研究生,研究方向:空间计量,经验似然。

1　主要结果和证明

记$A_n(\rho_1)=I_n-\rho_1 W_n, B_n(\rho_2)=I_n-\rho_2 M_n$并且假设$A_n(\rho_1)$和$B_n(\rho_2)$是非奇异矩阵。于是可以得到：

$$Y_n=A_n^{-1}(\rho_1)X_n\beta+A_n^{-1}(\rho_1)B_n^{-1}(\rho_2)\epsilon_{(n)}。$$

此时，假设$\epsilon_{(n)}$是正态分布的，

则Y_n服从期望为$A_n^{-1}(\rho_1)X_n\beta$，方差为$A_n^{-1}(\rho_1)B_n^{-1}(\rho_2)\sigma^2 I_n\left[A_n^{-1}(\rho_1)B_n^{-1}(\rho_2)\right]'$的正态分布。

于是，Y_n的拟似然函数为：

$$F=(2\pi)^{-\frac{n}{2}}\left|B_n(\rho_2)\right|\left|A_n(\rho_1)\right|\left(\sigma^{2n}\right)^{-\frac{1}{2}}\exp\left\{-\frac{1}{2\sigma^2}\epsilon_{(n)}'\epsilon_{(n)}\right\},$$

其中$\epsilon_{(n)}=B_n(\rho_2)\left\{A_n(\rho_1)Y_n-X_n\beta\right\}$。

进而，Y_n的拟对数似然函数为：

$$L=-\frac{n}{2}\log(2\pi)-\frac{n}{2}\log\sigma^2+\log\left|A_n(\rho_1)\right|+\log\left|B_n(\rho_2)\right|-\frac{1}{2\sigma^2}\epsilon_{(n)}'\epsilon_{(n)}$$

记$G_n=B_n(\rho_2)W_n A_n^{-1}(\rho_1)B_n^{-1}(\rho_2), H_n=M_n B_n^{-1}(\rho_2), \tilde{G}_n=\frac{1}{2}(G_n+G_n')$及$\tilde{H}_n=\frac{1}{2}(H_n+H_n')$。对对数似然函数求偏导可得：

$$\partial L/\partial\beta=\frac{1}{\sigma^2}X_n'B_n'(\rho_2)\epsilon_{(n)},$$

$$\partial L/\partial\rho_1=\frac{1}{\sigma^2}\left\{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n\beta\right\}^T\epsilon_{(n)}+\frac{1}{\sigma^2}\left\{\epsilon_{(n)}'\tilde{G}_n\epsilon_{(n)}-\sigma^2\operatorname{tr}(\tilde{G}_n)\right\},$$

$$\partial L/\partial\rho_2=\frac{1}{\sigma^2}\left\{\epsilon_{(n)}'\tilde{H}_n\epsilon_{(n)}-\sigma^2\operatorname{tr}(\tilde{H}_n)\right\},$$

$$\partial L/\partial\sigma^2=\frac{1}{2\sigma^4}\left(\epsilon_{(n)}'\epsilon_{(n)}-n\sigma^2\right)。$$

令上述偏导数等于0，可以获得以下估计方程：

$$X_n'B_n'(\rho_2)\epsilon_{(n)}=0, \tag{2}$$

$$\left\{B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n\beta\right\}'\epsilon_{(n)}+\left\{\epsilon_{(n)}'\tilde{G}_n\epsilon_{(n)}-\sigma^2\operatorname{tr}(\tilde{G}_n)\right\}=0, \tag{3}$$

$$\epsilon_{(n)}'\tilde{H}_n\epsilon_{(n)}-\sigma^2\operatorname{tr}(\tilde{H}_n)=0, \tag{4}$$

$$\epsilon_{(n)}'\epsilon_{(n)}-n\sigma^2=0。 \tag{5}$$

记$\tilde{g}_{ij}, \tilde{h}_{ij}, b_i$和$s_i$分别表示矩阵$\tilde{G}_n$第$i$行第$j$列的元素，矩阵$\tilde{H}_n$第$i$行第$j$列的元素，矩阵$X_n'B_n'(\rho_2)$第$i$列向量和向量$B_n(\rho_2)W_n A_n^{-1}(\rho_1)X_n\beta$的第$i$个元素。$\epsilon_i$表示$\epsilon_{(n)}=B_n(\rho_2)\left\{A_n(\rho_1)Y_n-X_n\beta\right\}$的第$i$个元素。并且规定，当求和的上标小于1时，我们令该和为0。为了处理(3)和(4)中的二次型形式，需要引入文献[5]中介绍的鞅差序列。定义σ—域：$\mathcal{F}_0=\{\Phi,\Omega\}, \mathcal{F}_i=\sigma(\epsilon_1,\epsilon_2,\cdots,\epsilon_i), 1\leqslant i\leqslant n$。令

$$\tilde{Y}_{in}=\tilde{g}_{ii}(\epsilon_i^2-\sigma^2)+2\epsilon_i\sum_{j=1}^{i-1}\tilde{g}_{ii}\epsilon_j, \tilde{Z}_{in}=\tilde{h}_{ii}(\epsilon_i^2-\sigma^2)+2\epsilon_i\sum_{j=1}^{i-1}\tilde{h}_{ii}\epsilon_j。 \tag{6}$$

则$\mathcal{F}_{i-1}\subseteq\mathcal{F}_i$，$\tilde{Y}_{in}$是$\mathcal{F}_i$—可测的，并且$E\left(\tilde{Y}_{in}|\mathcal{F}_{i-1}\right)=0$。因此，$\left\{\tilde{Y}_{in},\mathcal{F}_i,1\leqslant i\leqslant n\right\}$和$\left\{\tilde{Z}_{in},\mathcal{F}_i,1\leqslant i\leqslant n\right\}$构成两个鞅差序列，且

$$\epsilon_{(n)}'\tilde{G}_n\epsilon_{(n)}-\sigma^2\operatorname{tr}(\tilde{G}_n)=\sum_{i=1}^{n}\bar{Y}_{in}, \epsilon_{(n)}'\tilde{H}_n\epsilon_{(n)}-\sigma^2\operatorname{tr}(\tilde{H}_n)=\sum_{i=1}^{n}\bar{Z}_{in}。 \tag{7}$$

根据(2)到(7)可以得出参数$\theta\triangleq(\beta',\rho_1,\rho_2,\sigma^2)\in\mathbb{R}^{k+3}$的估计方程，令

$$\omega_i(\theta)=\begin{pmatrix} b_i\epsilon_i \\ \tilde{g}_{ii}(\epsilon_i^2-\sigma^2)+2\epsilon_i\sum_{j=1}^{i-1}\tilde{g}_{ij}\epsilon_j+s_i\epsilon_j \\ \tilde{h}_{ii}(\epsilon_i^2-\sigma^2)+2\epsilon_i\sum_{j=1}^{i-1}\tilde{h}_{ij}\epsilon_j \\ \epsilon_i^2-\sigma^2 \end{pmatrix}_{(k+3)\times 1}$$

通过$\omega_i(\theta)$提出θ的经验欧氏似然比统计量

$$L_n(\theta)=\sup_{p_i,1\leqslant i\leqslant n}\left\{-\frac{1}{2}\sum_{i=1}^{n}(np_i-1)^2\right\}$$

其中,$\{p_i\}$满足

$$p_i\geqslant 0,1\leqslant i\leqslant n,\sum_{i=1}^{n}p_i=1,$$

$$\sum_{i=1}^{n}p_ib_i\epsilon_i=0,$$

$$\sum_{i=1}^{n}p_i\left\{\tilde{g}_{ij}(\epsilon_i^2-\sigma^2)+2\epsilon_i\sum_{j=1}^{i-1}\tilde{g}_{ij}\epsilon_j+s_i\epsilon_j\right\}=0,$$

$$\sum_{i=1}^{n}p_i\left\{\tilde{h}_{ij}(\epsilon_i^2-\sigma^2)+2\epsilon_i\sum_{j=1}^{i-1}\tilde{h}_{ij}\epsilon_i\right\}=0,$$

$$\sum_{i=1}^{n}p_i(\epsilon_i^2-\sigma^2)=0。$$

在本文中,记$\mu_j=E(\epsilon_i^j)$,$j=3,4$,用$\mathrm{Vec}(\mathrm{diag}\,A)$表示由矩阵$A$的对角线上的元素构成的列向量,用$\|a\|$表示向量$a$的第二范数,用$1_n$表示由1作为元素组成的$n$维列向量,为了获得经验欧氏似然比统计量的渐近分布,需要给出以下假设条件:

A1.$\{\epsilon_i,1\leqslant i\leqslant n\}$是均值为0,方差有限的独立同分布随机变量序列,且存在$\eta_1>0$,使$E|\epsilon_1|^{4+\eta_1}<\infty$。

A2.假设$W_n,M_n,A_n^{-1}(\rho_1),B_n^{-1}(\rho_2)$及$\{x_i\}$满足以下条件:

(i)矩阵$W_n,M_n,A_n^{-1}(\rho_1),B_n^{-1}(\rho_2)$行元素和列元素的绝对值之和均一致有界;

(ii)$\{x_i\}$是一致有界的。

A3.存在常数$c_j>0$,$j=1,2$,使得$0<c_1\leqslant\lambda_{\min}(n^{-1}c)\leqslant\lambda_{\max}(n^{-1}\Sigma_{k+3})\leqslant c_2<\infty$,其中$\lambda_{\min}(A)$和$\lambda_{\max}(A)$分别表示矩阵$A$的最小和最大的特征值,且

$$\Sigma_{k+3}=\Sigma_{k+3}'=\mathrm{Cov}\left\{\sum_{i=1}^{n}\omega_i(\theta)\right\}=\begin{pmatrix}\Sigma_{11}&\Sigma_{12}&\Sigma_{13}&\Sigma_{14}\\ \Sigma_{21}&\Sigma_{22}&\Sigma_{23}&\Sigma_{24}\\ \Sigma_{31}&\Sigma_{32}&\Sigma_{33}&\Sigma_{34}\\ \Sigma_{41}&\Sigma_{42}&\Sigma_{43}&\Sigma_{44}\end{pmatrix},$$

其中,

$$\Sigma_{11}=\sigma^2\{B_n(\rho_2)X_n\}'B_n(\rho_2)X_n,$$

$$\Sigma_{12}=\sigma^2\{B_n(\rho_2)X_n\}'B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta+\mu_3\{B_n(\rho_2)X_n\}'\mathrm{Vec}(\mathrm{diag}\,\tilde{G}_n),$$

$$\Sigma_{13}=\mu_3\{B_n(\rho_2)X_n\}'\mathrm{Vec}(\mathrm{diag}\,\tilde{H}_n),$$

$$\Sigma_{14}=\mu_3\{B_n(\rho_2)X_n\}'I_n,$$

$$\Sigma_{22}=2\sigma^4\mathrm{tr}(\tilde{G}_n{}^2)+\sigma^2\{B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta\}'B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n+(\mu_4-3\sigma^4)\|\mathrm{Vec}(\mathrm{diag}\,\tilde{G}_n)\|^2+$$
$$2\mu_3\{B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta\}'\mathrm{Vec}(\mathrm{diag}\,\tilde{G}_n),$$

$$\Sigma_{23}=2\sigma^4\mathrm{tr}(\tilde{G}_n\tilde{H}_n)+(\mu_4-3\sigma^4)\mathrm{Vec}'(\mathrm{diag}\,\tilde{G}_n)\mathrm{Vec}(\mathrm{diag}\,\tilde{H}_n)+\mu_3\{B_n(\rho_2)W_nA_n^{-1}(\rho_1)X_n\beta\}'\mathrm{Vec}(\mathrm{diag}\,\tilde{H}_n),$$

$\Sigma_{24}=\left(\mu_4-\sigma^4\right)\operatorname{tr}\left(\tilde{G}_n\right)+\mu_3\left\{B_n\left(\rho_2\right)W_nA_n^{-1}\left(\rho_1\right)X_n\beta\right\}'I_n,$

$\Sigma_{33}=2\sigma^4\left(\tilde{H}_n{}^2\right)+\left(\mu_4-3\sigma^4\right)\|\operatorname{Vec}\left(\operatorname{diag}\tilde{H}_n\right)\|^2,$

$\Sigma_{34}=\left(\mu_4-\sigma^4\right)\operatorname{tr}\left(\tilde{H}_n\right),$

$\Sigma_{44}=n\left(\mu_4-\sigma^4\right)$。

引理 1　在(A1)～(A3)假设条件下，当$n\rightarrow\infty$时，有

$$\Sigma_{k+3}^{-1/2}\sum_{i=1}^{n}\omega_i(\theta)\xrightarrow{d}N\left(0,I_{k+3}\right)$$

$$n^{-1}\sum_{i=1}^{n}\omega_i(\theta)\omega_i'(\theta)=n^{-1}\Sigma_{k+3}+o_p(1)$$

见文献[4]的引理 3。

定理 1　在(A1)～(A3)假设条件下及模型(1)下，当$n\rightarrow\infty$时，有$-2L_n(\theta)\xrightarrow{d}\chi_{k+3}^2$，其中，$\chi_{k+3}^2$表示自由度为$k+3$的卡方分布。

证明：SARAR 模型下经验欧氏(对数)似然函数为：

$$L_n(\theta)=\sup_{p_i,1\leqslant i\leqslant n}\left\{-\frac{1}{2}\sum_{i=1}^{n}\left(np_i-1\right)^2\right\}$$

满足以下约束条件：① $\sum_{i=1}^{n}p_i=1$；② $p_i\geqslant 0,1\leqslant i\leqslant n$；③ $\sum_{i=1}^{n}p_i\omega_i(\theta)=0$。

利用拉格朗日乘子法给出$L_n(\theta)$的表达式。为此取

$$G=-\frac{1}{2}\sum_{i=1}^{n}\left(np_i-1\right)^2-nt'\sum_{i=1}^{n}p_i\omega_i(\theta)-\lambda\left(\sum_{i=1}^{n}p_i-1\right)\tag{8}$$

其中$t'\in\mathbb{R}^{k+3}$。对G求关于p_i的偏导数，得

$$\frac{\partial G}{\partial p_i}=-n\left(np_i-1\right)-nt'\omega_i(\theta)-\lambda$$

并令$\dfrac{\partial G}{\partial p_i}=0$，得

$$-n^2p_i+n-nt'\omega_i(\theta)-\lambda=0\tag{9}$$

在(9)式两边对i从 1 加到n求和，得

$$-n^2\sum_{i=1}^{n}p_i+\sum_{i=1}^{n}n-\sum_{i=1}^{n}nt'\omega_i(\theta)-\sum_{i=1}^{n}\lambda=0$$

利用约束条件① $\sum_{i=1}^{n}p_i=1$，得

$$\lambda=-\sum_{i=1}^{n}t'\omega_i(\theta)=-t'\sum_{i=1}^{n}\omega_i(\theta)=-nt'\bar{\omega}(\theta)\tag{10}$$

其中，$\bar{\omega}(\theta)=\dfrac{1}{n}\sum_{i=1}^{n}\omega_i(\theta)$。将(10)代入(9)中，得

$$-n^2p_i+n-nt'\omega_i(\theta)-\left(-nt'\bar{\omega}(\theta)\right)=0$$

上式化简可得

$$p_i=\frac{1}{n}+\frac{1}{n}t'\left[\bar{\omega}(\theta)-\omega_i(\theta)\right],1\leqslant i\leqslant n\tag{11}$$

将(11)代入约束条件③ $\sum_{i=1}^{n}p_i\omega_i(\theta)=0$，得

$$\begin{aligned}\sum_{i=1}^{n}p_i\omega_i(\theta)&=\sum_{i=1}^{n}\left\{\frac{1}{n}+\frac{1}{n}t'\left[\bar{\omega}(\theta)-\omega_i(\theta)\right]\right\}\omega_i(\theta)\\&=\sum_{i=1}^{n}\frac{1}{n}\omega_i(\theta)-\sum_{i=1}^{n}\frac{1}{n}t'\left[\omega_i(\theta)-\bar{\omega}(\theta)\right]\omega_i(\theta)\\&=\bar{\omega}(\theta)-\frac{1}{n}\sum_{i=1}^{n}t'\left[\omega_i(\theta)-\bar{\omega}(\theta)\right]\left[\omega_i(\theta)-\bar{\omega}(\theta)\right]\\&=\bar{\omega}(\theta)-\frac{1}{n}\sum_{i=1}^{n}\left[\omega_i(\theta)-\bar{\omega}(\theta)\right]\left[\omega_i(\theta)-\bar{\omega}(\theta)\right]'t\\&=\bar{\omega}(\theta)-S(\theta)t\end{aligned}$$

数学与计算机技术

其中 $S(\theta)=\frac{1}{n}\sum_{i=1}^{n}\left[\omega_i(\theta)-\overline{\omega}(\theta)\right]\left[\omega_i(\theta)-\overline{\omega}(\theta)\right]'$,将上式化简可得

$$t=S^{-1}(\theta)\overline{\omega}(\theta) \tag{12}$$

由(10)～(12)可得

$$\begin{aligned}-2L_n(\theta)&=\sum_{i=1}^{n}\left(np_i-1\right)^2\\&=\sum_{i=1}^{n}\left\{t'\left[\overline{\omega}(\theta)-\omega_i(\theta)\right]\right\}^2\\&=t'\sum_{i=1}^{n}\left[\overline{\omega}(\theta)-\omega_i(\theta)\right]\left[\overline{\omega}(\theta)-\omega_i(\theta)\right]'t\\&=nt'S(\theta)t\\&=n\left\{S^{-1}(\theta)\overline{\omega}(\theta)\right\}'S(\theta)S^{-1}(\theta)\overline{\omega}(\theta)\\&=n\overline{\omega}'(\theta)S^{-1}(\theta)\overline{\omega}(\theta)\end{aligned}$$

由引理1,得

$$-2L_n(\theta)\xrightarrow{d}\chi^2_{k+3}$$

由此完成了定理1的证明。

[参　考　文　献]

[1]OWEN A B. Empirical likelihood ratio confidence intervals for a single functional[J]. Biomertika,1988,75(2):237－249.

[2]QIN Y S. Empirical likelihood for Spatial Autoregressive Models with Spatial Autoregressive Disturbances[J]. The Indian Journal of Statistics,2021,Volume 83－A,Part1:1－25, https://doi.org/ 10.1007/s13171－019－00166－3.

[3]OWEN A B. Empirical likelihood for linear models[J]. Ann Statist,1991,19(4):1725－1747.

[4]罗旭.半参数模型的经验欧氏似然估计的大样本性质[J].应用概率统计,1994,10(4):344－352.

[5]KELEJIAN H H, PruchaI R. On the asymptotic distribution of the Moran I test statistic with applications[J].Journal of Econometrics,2001,104(2):219－257.

[责任编辑　苏　琴]

Empirical Euclidean Likelihood Inference for Spatial Autoregressive Model with Spatial Autoregressive Disturbances

TANG Jie

(College of Mathematics and Statistics, Guangxi Normal University, Guilin 541004, China)

Abstract: This paper studies the empirical euclidean likelihood inference for spatial autoregressive model with spatial autoregressive disturbances (SARAR model). Via empirical Euclidean likelihood methods and by means of a martingale difference array, we deal with the linear-quadratic : forms of the estimating equation of SARAR, construct the empirical Euclidean likelihood ratio statistics for the parameters in SARAR. It is shown that the empirical Euclidean likelihood ratio statistics are asymptotically chi-square distributed.

Keyword: SARAR model; martingale difference array; empirical Euclidean likelihood