第49卷第8期　　中国科学技术大学学报　　Vol. 49, No. 8
2019年8月　　JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA　　Aug. 2019

文章编号:0253-2778(2019)08-0630-05

空间自回归模型中参数的 Bayes 估计

吴世朋[1],张辉国[1],王合玲[2]

(1. 新疆大学数学与系统科学学院,新疆乌鲁木齐 830046;2. 新疆财经大学应用数学学院,新疆乌鲁木齐,830012)

摘要:首先采用线性 Bayes 方法估计了空间自回归模型的参数,并在均方误差矩阵准则下研究了线性 Bayes 估计相对两步最小二乘估计的优良性. 然后,使用 Metropolis 抽样算法实现了对空间自相关系数的估计. 最后,通过模拟试验比较了线性 Bayes 估计与两步最小二乘估计的优缺点.

关键词:空间自回归模型;线性 Bayes 估计;两步最小二乘估计

中图分类号:O212.1　　**文献标识码**:A　　doi:10.3969/j.issn.0253-2778.2019.08.006

2010 Mathematics Subject Classification:62C10;62F15;62J05

引用格式:吴世朋,张辉国,王合玲. 空间自回归模型中参数的 Bayes 估计[J]. 中国科学技术大学学报,2019,49(8):630-634.

WU Shipeng, ZHANG Huiguo, WANG Heling. The Bayes estimation of parameters of spatial autoregressive model[J]. Journal of University of Science and Technology of China, 2019,49(8):630-634.

The Bayes estimation of parameters of spatial autoregressive model

WU Shipeng[1], ZHANG Huiguo[1], WANG Heling[2]

(1. *College of Mathematics and System Science, Xinjiang University, Urumqi* 830046, *China*;
2. *College of Applied Mathematics, Xinjiang University of Finance & Economics, Urumqi* 830012, *China*)

Abstract: First, the linear Bayes was used to estimate the parameters of spatial autoregressive model, and the superiorities of the linear Bayes estimator over two-step least square estimator were studied in terms of the mean square error matrix (MSEM) criterion. Then, the estimation of spatial autocorrelation coefficient was implemented by Metropolis algorithm. Finally, the superiority of the linear Bayes estimation and two-step least square estimation was compared by simulation experiments.

Key words: spatial autoregressive model; linear Bayes estimation; two-step least square estimation

0　引言

空间自回归模型(spatial autoregressive model,SAR)是空间计量经济学中最重要的模型之一,受到了统计学与经济学等领域众多学者的关注. 模型的估计方法在近几十年来不断涌现,Anselin[1] 将该模型称为“混合回归-空间自回归”模型并给出了其最大似然估计;Lesage[2] 使用 Bayesian 方法对该模型进行了估计;Lee [3] 使用两阶段最小二乘估计(two-stage least square, 2SLS)和 Kelejian 等[4] 采用工具变量法分别估计了模型的参数. 在许多实

收稿日期:2018-10-23;**修回日期**:2019-04-10

基金项目:新疆自然科学基金(2019D01C045),国家社会科学基金项目(16BTJ024),教育部人文社会科学研究规划基金项目(19YJA910007)资助.

作者简介:吴世朋,男,1993 年生,硕士生. 研究方向:贝叶斯空间计量模型. E-mail:wushipeng1234@163.com

通讯作者:张辉国,博士/副教授. E-mail:zhanghg@xju.edu.cn

际问题中,人们通常缺乏对模型参数和误差项的独立性及分布的精确认识,只知道它们一、二阶矩方面的统计特性,这时线性 Bayes 估计较为适合此类问题的解决.目前,已有多篇文献[5-7]讨论了该方法对线性模型的估计以及优良性问题,但尚未发现对 SAR 模型的线性 Bayes 估计.基于此,本文给出了 SAR 模型的线性 Bayes 估计并证明了估计量在均方误差矩阵(mean square error matrix,MSEM)准则下的优良性.最后,通过编程进行模拟试验,对两种方法的特点进行了讨论.模型形式如下所示:

$$\left.\begin{aligned} &Y=\rho WY+X\beta+\varepsilon,\\ &E(\varepsilon)=0,\mathrm{Cov}(\varepsilon)=\sigma^2 I_p \end{aligned}\right\} \tag{1}$$

式中,Y 是 $n\times 1$ 维的观测向量,X 是 $n\times p$ 维的设计矩阵,ρ 是空间自相关系数,β 是 $p\times 1$ 维的待估参数,W 是空间权重矩阵(spatial weights matrix),一般依据观测点相邻关系或位置关系确定,有关的详细介绍可参考文献[1].

为了说明线性 Bayes 估计的优良性,首先采用文献[8,15]中给出的两步最小二乘法对 SAR 模型进行参数估计.

对于模型(1),假设参数 ρ 已知,则该模型转化为普通的线性回归模型

$$Y^*=X\beta+\varepsilon \tag{2}$$

式中,$Y^*=Y-\rho WY$.对于该模型,利用最小二乘估计方法可得

$$\hat{\beta}_{\mathrm{2LS}}=(X'X)^{-1}X'Y^* \tag{3}$$

进而可以得到

$$\hat{\beta}_{\mathrm{2LS}}=(X'X)^{-1}X'(Y-\rho WY) \tag{4}$$

将 $\hat{\beta}_{\mathrm{2LS}}$ 代入模型(1)中整理得

$$(I-S)Y=\rho(I-S)WY+\varepsilon \tag{5}$$

式中,$S=X(X'X)^{-1}X'$.因此,可得 ρ 的最小二乘估计为

$$\hat{\rho}_{\mathrm{2LS}}=[Y'W'(I-S)WY]^{-1}Y'W'(I-S)Y \tag{6}$$

值得注意的是,模型(5)为典型的一阶空间自回归模型,Anselin[9]证明了该模型参数 ρ 的最小二乘估计是有偏的且不具备相合性,李序颖和顾岚[10]对该结果重新进行了证明.因此,采用上述两步最小二乘方法估计出来的 $\hat{\rho}_{\mathrm{2LS}}$ 也不具有无偏性和相合性,本文的模拟试验也证实该证明结果.

本文内容安排如下:节 1 给出了 SAR 模型的 Bayes 估计,节 2 讨论了在 MSEM 准则下 $\hat{\beta}_{\mathrm{BE}}$ 相对于 $\hat{\beta}_{\mathrm{2LS}}$ 的优良性,节 3 采用 Monte Carlo 方法进行模拟试验,讨论了 $\hat{\beta}_{\mathrm{BE}}$,$\hat{\beta}_{\mathrm{2LS}}$,$\hat{\rho}_{\mathrm{BE}}$,$\hat{\rho}_{\mathrm{2LS}}$ 四个估计量的数值模拟结果.

1　Bayes 线性最小风险估计

模型参数的贝叶斯估计方法大致可以分为两种:一种是假定模型参数的先验分布为共轭先验,如正态分布或无信息先验分布,则在二次损失下 Bayes 估计由后验均值给出,文献[2]正是基于该方法得出了 SAR 模型的 Bayes 估计;另外一种是在 Gauss-Markov 模型下,通过假定参数与误差项的二阶矩存在,在与 Y 相关的类中寻找参数的最优估计量,此估计通常称为线性 Bayes 估计.后者最早由 Rao[11]提出,Gruber[12],Wei 和 Zhang 等[5-7,13]都曾使用该方法对线性模型的参数进行估计并讨论其优良性的问题.本文借助该方法对空间自回归模型进行了研究.

设 β,ρ 的先验分布分别满足下列条件:

$$\left.\begin{aligned} &\pi(\beta):E(\beta)=\mu,\mathrm{Cov}(\beta)=\tau^2 I_p,\\ &\pi(\rho):E(\rho)=\lambda \end{aligned}\right\} \tag{7}$$

式中,μ 是已知的向量,λ 是已知的标量.

为了进行 β 的线性 Bayes 估计,首先需要定义一个损失函数:

$$L(\delta,\beta)=(\delta-\beta)'D(\delta-\beta) \tag{8}$$

式中,D 为已知的正定矩阵,进而定义相应的风险函数:

$$R(\delta,\beta)=E[L(\delta,\beta)] \tag{9}$$

式中,E 表示对联合分布 (Y,β) 求期望.

首先假设 ρ 已知,则可设模型(2)中参数向量 β 的线性无偏估计类:

$$F=\left\{\begin{aligned} &\hat{\beta}^*=AY^*+b:A\text{ 是 }p\times n\text{ 维矩阵},\\ &b\text{ 是 }p\times 1\text{ 维向量} \end{aligned}\right\}.$$

记 β 的 Bayes 线性无偏估计(Bayes linear unbiased estimator,BLUE)为 $\hat{\beta}_{\mathrm{BE}}$,按定义 $\hat{\beta}_{\mathrm{BE}}$ 需满足:

$$\begin{aligned} R(\hat{\beta}_{\mathrm{BE}},\beta)=\min_{A,b}R(\hat{\beta}^*,\beta)=\\ \min_{A,b}E[(\hat{\beta}^*-\beta)'D(\hat{\beta}^*-\beta)] \end{aligned}$$

和无偏性条件 $E(\hat{\beta}^*-\beta)=0$.解无偏性条件可得

$$b=(I_p-AX)\mu \tag{10}$$

为了求解出矩阵 A,需要首先计算 β 的 Bayes 风险,令 $\dfrac{\partial R(\hat{\beta}^*,\beta)}{\partial A}=0$,由矩阵微商法则可得

$$A=k^{-1}X'(I_n+k^{-1}XX')^{-1} \tag{11}$$

式中，$k=\sigma^2/\tau^2$. 由矩阵逆运算 $(P+BCB')^{-1}=P^{-1}-P^{-1}B(B'P^{-1}B+C^{-1})^{-1}B'P^{-1}$ 可得

$$A=(X'X+kI_p)^{-1}X' \tag{12}$$

因此，β 的 BLUE 为

$$\begin{aligned}\hat{\beta}_{\mathrm{BE}}=(X'X+kI_p)^{-1}(X'Y^*+k\mu)=\\(X'X+kI_p)^{-1}[X'(Y-\rho WY)+k\mu]=\\(X'X+kI_p)^{-1}(X'X\hat{\beta}_{2\mathrm{LS}}+k\mu)\end{aligned} \tag{13}$$

对于空间效应参数 ρ 的估计，本文采用的是 Metropolis-Gibbs 算法[14]. 首先，假设参数 (ρ,σ) 的先验分布信息如式(14)所示.

$$\pi(\rho)\propto\lambda,\pi(\sigma)\propto(1/\sigma) \tag{14}$$

由贝叶斯定理可得参数 $p(\rho,\sigma\mid y)$ 的联合后验分布为

$$\begin{aligned}p(\rho,\sigma,\beta\mid y)\propto\mid I_n-\rho W\mid\sigma^{-(n+1)}\cdot\\\exp\{-\frac{1}{2\sigma^2}(y-X\beta-\rho Wy)'(y-X\beta-\rho Wy)\}\end{aligned} \tag{15}$$

通过式(13)得出了 β 的线性 Bayes 无偏估计 $\hat{\beta}_{\mathrm{BE}}$，下面考虑参数 ρ 和 σ 的估计. 若假设 ρ 为已知参数，那么参数 σ 的条件后验为

$$p(\sigma\mid y)\propto\sigma^{-(n+1)}\exp\{-\frac{1}{2\sigma^2}\varepsilon'\varepsilon\} \tag{16}$$

式中，$\varepsilon=y-X\beta-\rho Wy$. 由式(16)的形式可以推知 σ^2 服从 $\chi^2(n)$ 分布，参数的分布为常见的已知形式，故可采用 Gibbs 抽样方法.

在 $\sigma^2\sim\chi^2(n)$ 的条件下，可以得出 ρ 的后验分布形式为

$$\begin{aligned}p(\rho\mid\sigma,y,\beta)\propto\sigma^{-n/2}\mid I_n-\rho W\mid\cdot\\\{(y-X\beta-\rho Wy)'(y-X\beta-\rho Wy)\}^{-n/2}\end{aligned} \tag{17}$$

由式(17)可知，ρ 的后验分布形式不再是常见分布形式，适合采用 Metropolis 抽样法.

对于 Metropolis 抽样，采用 $N(0,1)$ 作为建议分布，并结合当前值 ρ^c 产生参数 ρ 的备选值 ρ^*，如式(18)所示：

$$\rho^*=\rho^c+c\cdot N(0,1) \tag{18}$$

式中，c 被称作调和参数. 将备选值 ρ^* 与当前值 ρ^c 代入到式(19)中计算接受概率

$$q(\rho^c,\rho^*)=\min[1,\frac{p(\rho^*\mid\sigma,\beta,y)}{p(\rho^c\mid\sigma,\beta,y)}] \tag{19}$$

最后，将产生的参数 ρ 的序列平均值作为估计值.

2 MSEM 准则下 BLMRE 的优良性

为了研究 $\hat{\beta}_{\mathrm{BE}}$ 在 MSEM 准则下的优良性，首先需要引入 MSEM 的定义.

定义 2.1 参数向量 θ 的估计为 $\hat{\theta}$，则 $\hat{\theta}$ 的均方误差(MSE)定义为

$$\mathrm{MSE}(\theta)=E[(\hat{\theta}-\theta)'(\hat{\theta}-\theta)].$$

$\hat{\theta}$ 的 MSEM 的定义为

$$M(\theta)=E[(\hat{\theta}-\theta)(\hat{\theta}-\theta)'].$$

设参数向量 θ 两种不用方法的估计分别为 $\hat{\theta}_1$ 和 $\hat{\theta}_2$，若

$$M(\hat{\theta}_2)-M(\hat{\theta}_1)\geqslant 0$$

$$(\text{或 }\mathrm{MSE}(\hat{\theta}_2)-\mathrm{MSE}(\hat{\theta}_1)\geqslant 0),$$

则称 $\hat{\theta}_1$ 在 MSEM(或 MSE)准则下优于 $\hat{\theta}_2$. 显然，MSEM 准则强于 MSE 准则.

定理 2.1 模型(1)中参数 β 的两步最小二乘估计(2LS)和 Bayes 线性无偏估计(BLUE)分别由式(4)和式(13)给出，则

$$M(\hat{\beta}_{2\mathrm{LS}})-M(\hat{\beta}_{\mathrm{BE}})>0.$$

证明 由定义 2.1 和式(4)可知

$$\begin{aligned}M(\hat{\beta}_{2\mathrm{LS}})=E[(\hat{\beta}_{2\mathrm{LS}}-\beta)(\hat{\beta}_{2\mathrm{LS}}-\beta)']=\\E\{E[(\hat{\beta}_{2\mathrm{LS}}-\beta)(\hat{\beta}_{2\mathrm{LS}}-\beta)']\mid\beta\}=\\E[(X'X)^{-1}X'Y^*-\beta][(X'X)^{-1}X'Y^*-\beta]'=\\E[(X'X)^{-1}X'(X\beta+\varepsilon)-\beta]\cdot\\[(X'X)^{-1}X'(X\beta+\varepsilon)-\beta]'=\\E[(X'X)^{-1}X'\varepsilon][(X'X)^{-1}X'\varepsilon]'=\\\sigma^2(X'X)^{-1}.\end{aligned} \tag{20}$$

由定义 2.1 和式(13)可知

$$\begin{aligned}M(\hat{\beta}_{\mathrm{BE}})=E[(\hat{\beta}_{\mathrm{BE}}-\beta)(\hat{\beta}_{\mathrm{BE}}-\beta)']=\\\mathrm{Cov}(\hat{\beta}_{\mathrm{BE}}-\beta)=\\E[\mathrm{Cov}(\hat{\beta}_{\mathrm{BE}}-\beta\mid\beta)]+\mathrm{Cov}[E(\hat{\beta}_{\mathrm{BE}}-\beta\mid\beta)]=\\\sigma^2(X'X+kI_p)^{-1}X'X(X'X+kI_p)^{-1}+\\\tau^2k^2(X'X+kI_p)^{-2}=\\\sigma^2(X'X+kI_p)^{-1}\end{aligned} \tag{21}$$

结合式(20)和式(21)得

$$\begin{aligned}M(\hat{\beta}_{2\mathrm{LS}})-M(\hat{\beta}_{\mathrm{BE}})=\\\sigma^2[(X'X)^{-1}-(X'X+kI_p)^{-1}]>0\end{aligned} \tag{22}$$

由定理 2.1，利用公式 $\mathrm{MSE}(\hat{\theta})=\mathrm{tr}(M(\hat{\theta}))$，可以得出下列推论.

推论 2.1 设 SAR 模型中参数 β 的两步最小二乘估计和线性 Bayes 无偏估计分别由式(4)和式(13)给出，则

$$\mathrm{MSE}(\hat{\beta}_{2\mathrm{LS}})-\mathrm{MSE}(\hat{\beta}_{\mathrm{BE}})>0.$$

3　Monte Carlo 模拟实验

LeSage[15]基于 MATLAB 编制出的空间计量工具箱，其中已经包含了空间自回归模型的参数估计，但该方法是通过假定参数向量的先验分布为正态分布或者无信息分布. 尚未有基于线性 Bayes 方法估计空间自回归模型. 本文使用 Anselin 的哥伦布社区犯罪数据集[14]中的空间邻接矩阵进行模拟试验，讨论本文中的两步最小二乘估计和线性 Bayes 估计两种方法的优缺点. 为了避免单次试验的偶然性，文中的所有试验均进行 $N=100$ 次. 试验结果如表 1 所示.

表 1　100 次模拟试验参数估计结果

Tab. 1　Estimation results for the parameters under $N=100$ replications

参数	真值	两步最小二乘估计			线性 Bayes 估计		
		均值	方差	中值	均值	方差	中值
ρ	0.3	0.304 3	0.032 9	0.300 4	0.300 4	0.010 5	0.299 2
β_1	1	0.958 0	0.260 4	0.968 1	0.974 9	0.214 7	0.995 0
β_2	3	2.993 0	0.218 5	3.005 3	3.002 9	0.218 4	3.025 9
σ^2	0.36	0.122 3	0.024 7	0.119 1	0.250 3	0.214 8	0.185 5

由表 1 可知，上述两种方法在参数估计的精确度方面均取得了较好的结果，相比较而言，线性 Bayes 估计得到的结果更为精确，这是因为线性 Bayes 估计利用了参数的先验矩信息. 另一方面，两步最小二乘估计的推导过程相对简单，程序的编程也相对简单. 在模型计算的用时方面，两步最小二乘估计的用时大约为 0.066 s，而线性 Bayes 估计用时较长，大约为 31.141 s，这是因为线性 Bayes 估计在估计参数 ρ 时采用的是 Metropolis-Gibbs 算法，涉及的运算量较大，导致耗时较多.

线性 Bayes 估计中参数 ρ 的估计用到了 Metropolis 抽样，需要检测生成的 MCMC 链是否达到了收敛. 本文采用了 MATLAB 中的函数 coda 对生成的 MCMC 链进行收敛诊断，包括自相关性检验、MCMC 收敛诊断[16]和 χ^2 收敛检验[17]. 参数 ρ 的 MCMC 收敛诊断结果见表 2，可知关于 ρ 的 MCMC 链已经达到了收敛的状态.

表 2　参数 ρ 的 MCMC 收敛诊断结果

Tab. 2　MCMC convergence diagnostics for parameter ρ

自相关性	Lag1	Lag5	Lag10	Lag50	
	0.770	0.309	0.165	−0.069	
Raftery-Lewis 检验	$q=0.025$	$r=0.010$	$s=0.950$		
	Thin	Burn	Total	Nmin	I-stat
	1	32	8906	937	9.505
Geweke 卡方检验（样本的前 20% v. s. 后 50%）	NSE 估计	均值	N. S. E.	概率	
	i. i. d.	0.2911	0.0003	0.2159	
	4% taper	0.2912	0.0007	0.6015	
	8% taper	0.2912	0.0007	0.6268	
	15% taper	0.2911	0.0007	0.5954	

4　结论

本文采用了线性 Bayes 估计对空间自回归模型进行了研究，并在均方误差矩阵准则下证明了线性 Bayes 估计相对于两步最小估计的优良性. 该方法丰富了空间自回归模型参数的估计的方法库，编译的 R 程序（可通过联系作者获取）也便于采用空间自回归模型进行实证研究，这也将进一步促进该类模型的推广运用.

通过模拟实验发现两种方法存在以下两个问

题:第一,当模型中包含异常值或者误差项由 t 分布或者柯西分布中抽取时,无论是两步最小二乘估计方法还是线性 Bayes 估计,其结果并不是很理想,模型参数的估计值与真实值存在较大的偏差. 因此,可以借鉴 Lesage[2] 对空间自回归模型中误差项的处理方法,将误差项设置成为异方差形式,并给予相应的先验分布,从而达到能够对包含异常值的数据进行有效估计,这也是本文下一步的研究方向;第二,本文采用线性 Bayes 估计进行模型参数的计算时,求 ρ 所对应的 $I-\rho W$ 的行列式值并取对数,那么当 $|I-\rho W|$ 取到负值的时候,该项计算便不能进行,而两步估计避免了计算 $|I-\rho W|$. 因此,在邻接权重矩阵 W 取特殊值的情况下,选择两步法进行估计成为现有唯一的选择,有必要针对这一状况进行进一步的研究.

参考文献(References)

[1] ANSELIN L. Spatial Econometrics: Methods and Models[M]. Netherlands: Springer, 1988.

[2] LESAGE J P. Bayesian estimation of spatial autoregressive models [J]. International Regional Science Review, 1997, 20(1-2): 113-129.

[3] LEE L F. Best spatial two-stage least squares estimators for a spatial autoregressive model with autoregressive disturbances[J]. Econometric Reviews, 2003, 22(4): 307-335.

[4] KELEJIAN H H, PRUCHA I R, YUZEFOVICH Y. Instrumental variable estimation of a spatial autoregressive model with autoregressive disturbances: Large and small sample results [J]. Advances in Econometrics, 2004, 18(18): 163-198.

[5] WEI Laisheng, ZHANG Weiping. The superiorities of Bayes linear minimum risk estimation in linear model [J]. Communications in Statistics, 2007, 36(5): 917-926.

[6] TRENKLER G, WEI L. The Bayes estimator in a misspecified linear regression model[J]. Test, 1996, 5(1): 113-123.

[7] 霍涉云,张伟平,韦来生. 一类线性模型参数的 Bayes 估计及其优良性[J]. 中国科学技术大学学报,2007,37(7):773-776.
HUO Sheyun, ZHANG Weiping, WEI Laisheng. The superiorities of Bayes estimation in a class of linear model [J]. Journal of University of Science and Technology of China, 2007, 37(7): 773- 776.

[8] 魏传华,梅长林. 半参数空间变系数回归模型的两步估计方法及其数值模拟[J]. 统计与信息论坛,2005,20(1):16-19.
WEI Chuanhua, MEI Changlin. Two-step procedure and numerical simulations for semiparametric spatially varying-coefficient regression model[J]. Statistics & Information Forum, 2005, 20(1): 16-19.

[9] 魏传华,胡晶,吴喜之. 空间自相关地理加权回归模型的估计[J]. 数学的实践与认识,2010,40(22):126-134.
WEI Chuanhua, HU Jing, WU Xizhi. Estimation in geographically weighted regression with spatial autocorrelation [J]. Mathematics in Practice and Theory, 2010, 40(22): 126-134.

[10] LI Xuying, GU Lan. The spatial autoregressive model and its estimation[J]. Statistical Research, 2004, 21(6): 48-51.
李序颖,顾岚. 空间自回归模型及其估计[J]. 统计研究,2004,21(6): 48-51.

[11] RAO C R. Linear Statistical Inference and Its Applications[M]. New York: Wiley, 1973.

[12] GRUBER M H J. Regression Estimators: A Comparative Study[M]. Boston: Academic Press, 1990.

[13] 韦来生. 错误先验假定下回归系数 Bayes 估计的小样本性质[J]. 应用概率统计,2000,16(1): 71- 80.
WEI Laisheng. The small-sample properties for the Bayes estimator of regression coefficients under misspecified prior assumption[J]. Chinese Journal of Applied Probability and Statistics, 2000, 16(1): 71-80.

[14] LESAGE J P. The Theory and Practice of Spatial Econometrics [R]. Toledo, OH: University of Toledo, 1999.

[15] LESAGE J P. Applied Econometrics Using MATLAB [R]. Toledo, OH: University of Toledo, 1999.

[16] RAFTERY A E, LEWIS S M. The number of iterations, convergence diagnostics and generic metropolis algorithms in practical Markov chain Monte Carlo[C]// Practical Markov Chain Monte Carlo. London: Chapman and Hall, 1995.

[17] GEWEKE J. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments [C]// Bayesian Statistics. Oxford: Clarendon Press, 1992: 169-193.