

研究生阶段汇报材料

唐 洁

(广西师范大学数学与统计学院, 广西桂林, 541004)

AEL apply to SARAR model

Tang Jie

2021 年 11 月 12 日

摘要

这是一篇关于调整经验似然应用于 SARAR 模型的论文进度报告。

第一次汇报，单参数情形下，小样本情形下 AEL 明显好于 EL。

第二次汇报，多参数情形下，小样本情形下 AEL 明显好于 EL。

第三次汇报，寻找新算法使得 λ 满足 AEL 约束条件。

第四次汇报，新算法下，不仅 λ 满足 AEL 约束条件，还证明了多参数情形下，AEL 效果好于 EL，尤其在小样本情形下。并应用到 SARAR 模型中。

同时提到了接下来的任务，一是复刻 Chen431 页图 1，二是复刻 Chen435 页表 2 的数据，三是实现 Chen433 页算法，四是理论证明。

第五次汇报，复刻 Chen431 页图 1；阅读文献并整理部分证明。

目录

1	第一次汇报	1
1.1	单参数情形下：AEL 比 EL 好	1
1.1.1	标准正态样本-单参数-小样本	1
1.1.2	标准正态样本-单参数-大样本	1
2	第二次汇报	2
2.1	多参数情形下：AEL 比 EL 好	2
3	第三次汇报	3
3.1	发现问题	3
3.2	解决问题	3
3.2.1	Chen 算法	3
4	第四次汇报	4
4.1	多参数情形下：AEL 比 EL 好	4
4.1.1	线性模型	4
4.1.2	SARAR 模型	5
5	第五次汇报	7
6	题外话	8

1 第一次汇报

1.1 单参数情形下：AEL 比 EL 好

对 n 个来自标准正态分布的同一个样本，分别进行 EL 与 AEL 数值模拟，并进行重复 $nsim$ 次。

1.1.1 标准正态样本-单参数-小样本

表 1: Coverage probabilities of population mean

Normal data								
	n=10				n=20			
NV	0.80	0.90	0.95	0.99	0.80	0.90	0.95	0.99
EL	?				0.7768	0.8790	0.9282	0.9816
AEL	0.8	0.896	0.939	0.999	0.8140	0.9076	0.9452	0.9906

NV = nominal value; EL = empirical likelihood; AEL = Adjusted EL.

小样本情形下，当 $n=10$ 时，EL 出现无法计算的情况。与此相比，AEL 一定有解，并且仍有不错的结果。当 $n=20$ 时，分别在名义水平 0.8、0.9、0.95、0.99 情形下，明显可看出 AEL 比 EL 效果更好。

1.1.2 标准正态样本-单参数-大样本

表 2: Coverage probabilities of population mean

Normal data								
	n=100				n=500			
NV	0.80	0.90	0.95	0.99	0.80	0.90	0.95	0.99
EL	0.7926	0.892	0.9494	0.9904	0.8024	0.9026	0.9508	0.9884
AEL	0.8042	0.9012	0.9556	0.9918	0.8064	0.9044	0.9520	0.9890

NV = nominal value; EL = empirical likelihood; AEL = Adjusted EL.

大样本情形下，AEL 与 EL 趋于一致，AEL 仍比 EL 的覆盖率高一些。
结论：单参数情形下，无论大样本还是小样本，AEL 都比 EL 好。

2 第二次汇报

2.1 多参数情形下：AEL 比 EL 好

重新计算，结果见第四次汇报。

3 第三次汇报

3.1 发现问题

牛顿迭代法求解 $G(\lambda) = 0$ 的程序，在经验似然下， λ 计算上满足 $G(\lambda)=0$ 的约束条件，而在调整经验似然下，求解出的 λ 都没有满足 $G(\lambda)=0$ 的约束条件。于是寻找新的算法，能在调整经验方法下，求得满足所有约束下的 λ 的值。

3.2 解决问题

踏上寻“根”之路。

此根须同时满足 (A1) $G(\lambda) = 0$; (A2) $1 + \lambda' g_i > 0, i = 1, 2, \dots, n, n+1$ 。

3.2.1 Chen 算法

Chen 在 433 页给了他的算法，那就试试看吧。

$$\begin{aligned} a_n &= \max(1, \log(n)/2) \\ g_{n+1} &= -a_n \bar{g}_n \\ R(\lambda) &= \sum_{i=1}^{n+1} \log(1 + \lambda' g_i) \end{aligned} \tag{1}$$

开始 Chen 的程序之前，先分别计算 $R(\lambda)$ 的一阶偏导数 \dot{R} 和二阶偏导数 \ddot{R} 。

$$\begin{aligned} \dot{R} &= \sum_{i=1}^{n+1} \frac{g_i}{1 + \lambda' g_i} \\ \ddot{R} &= \sum_{i=1}^{n+1} -\frac{g_i g_i'}{(1 + \lambda' g_i)^2} \end{aligned} \tag{2}$$

最后经过部分调整，Chen 算法落地成功。

Chen 算法不论在经验似然还是调整经验似然下，求解的 λ 都满足约束要求， λ 百分之百满足 $|G(\lambda)| \leq 10^{-6}$ 。

4 第四次汇报

4.1 多参数情形下：AEL 比 EL 好

来自多参数模型的 n 个样本，名义水平为 0.95，分别计算 EL 与 AEL 方法下的覆盖率，并进行重复 2000 次。

4.1.1 线性模型

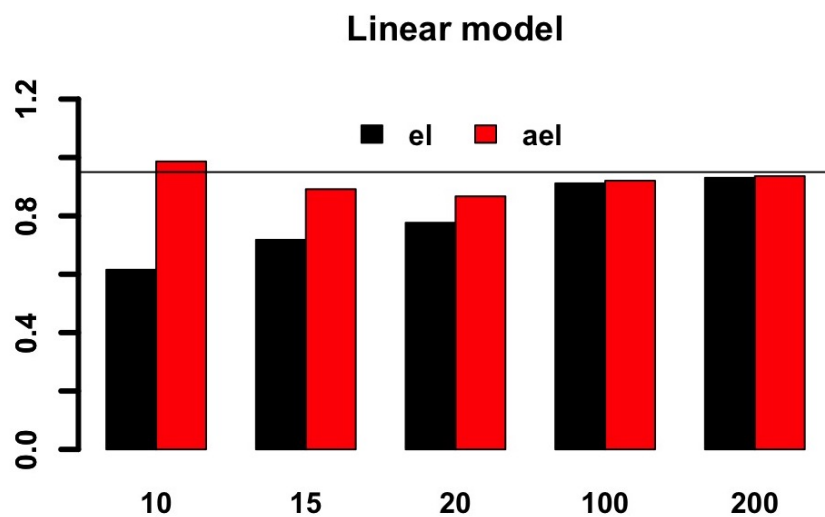
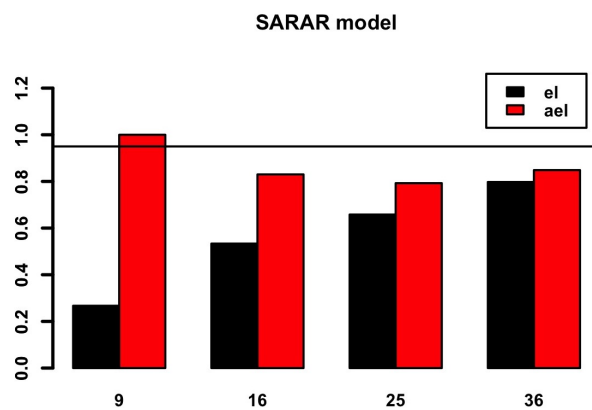


图 1: 线性模型 EL 与 AEL 对比图

由图可见，小样本 ($n=10, 15, 20$) 的表现在 AEL 方法下显著提高，在大样本 ($n=100, 200$) 下，EL 与 AEL 趋于相近。

结论：线性模型下，无论大样本还是小样本，AEL 都比 EL 好。

4.1.2 SARAR 模型

图 2: W_n 分别取 $grid_9$ 、 $grid_{16}$ 、 $grid_{25}$ 、 $grid_{36}$

模拟 2000 次, 置信水平为 0.95, $\beta=3.5$, $(\rho_1, \rho_2)=(-0.85, -0.15)$, X_n 由标准正态数据的平方生成时, 并分别在样本为 9、16、25、36 个的情形下计算覆盖率。由图可知, 小样本情形下, AEL 明显比 EL 效果更好。

图 3: W_n 分别取 $grid_{49}$ 、 $grid_{100}$ 、 $grid_{169}$

参数同上, 分别在样本为 49、100、169 个的情形下计算覆盖率。图 6 可看出, 大样本情形下, AEL 与 EL 很接近, 但 AEL 仍比 EL 略好一些。

结论: SARAR 模型下, 无论大样本还是小样本, AEL 都比 EL 好。

表 3: Coverage probabilities of the EL and AEL confidence regions with $\epsilon_i \sim N(0, 1)$

(ρ_1, ρ_2)	Wn=Mn	EL	AEL	(ρ_1, ρ_2)	Wn=Mn	EL	AEL
(-0.85, -0.15)	<i>grid</i> ₉	0.2670	1	(-0.85, 0.15)	<i>grid</i> ₉	0.2800	1
	<i>grid</i> ₁₆	0.5335	0.8300		<i>grid</i> ₁₆	0.5160	0.8385
	<i>grid</i> ₂₅	0.6580	0.7925		<i>grid</i> ₂₅	0.6910	0.7925
	<i>grid</i> ₃₆	0.7975	0.8485		<i>grid</i> ₃₆	0.7515	0.8105
	<i>grid</i> ₄₉	0.8100	0.8515		<i>grid</i> ₄₉	0.8400	0.8655
	<i>grid</i> ₁₀₀	0.8915	0.9040		<i>grid</i> ₁₀₀	0.899	0.9110
	<i>grid</i> ₁₆₉	0.9210	0.9340		<i>grid</i> ₁₆₉	0.925	0.9325
(0.85, -0.15)	<i>grid</i> ₉	0.2470	1	(0.85, 0.15)	<i>grid</i> ₉	0.2440	1
	<i>grid</i> ₁₆	0.5310	0.813		<i>grid</i> ₁₆	0.4795	0.8120
	<i>grid</i> ₂₅	0.6805	0.7895		<i>grid</i> ₂₅	0.6810	0.7815
	<i>grid</i> ₃₆	0.7710	0.8310		<i>grid</i> ₃₆	0.7430	0.8075
	<i>grid</i> ₄₉	0.8305	0.8530		<i>grid</i> ₄₉	0.8130	0.8505
	<i>grid</i> ₁₀₀	0.9070	0.9170		<i>grid</i> ₁₀₀	0.9070	0.9205
	<i>grid</i> ₁₆₉	0.9240	0.9310		<i>grid</i> ₁₆₉	0.9270	0.9325

5 第五次汇报

来自独立的标准二维正态分布 50 个观测值, 截面似然于真值 $(\mu, \sigma^2) = (0, 1)$, 可以画出 $g_i, i = 1, 2, \dots, n, n+1$ 值。

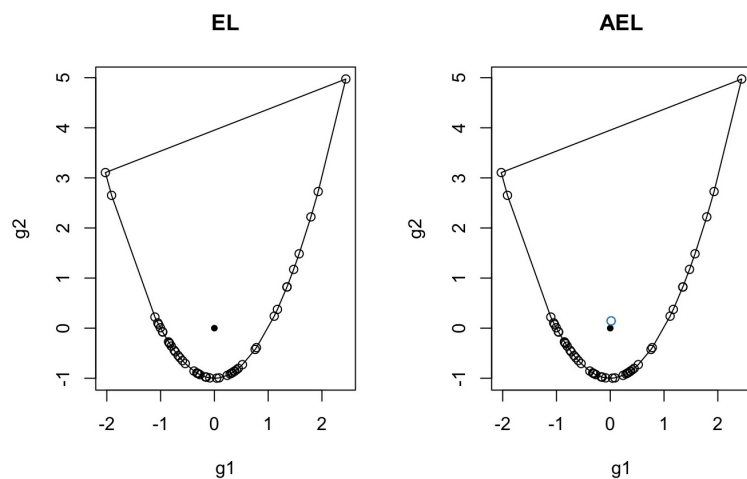


图 4: 真值: 凸包 (左) 和调整凸包 (右), 黑点是 (0, 0).

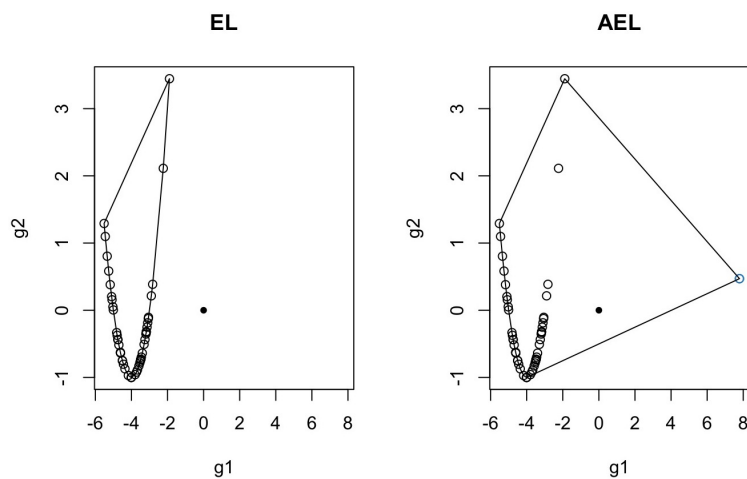


图 5: 非真值: 凸包 (左) 和调整凸包 (右), 黑点是 (0, 0).

截面似然于非真值 $(\mu, \sigma^2) = (4, 1)$ 。

6 题外话

非常感谢老师的耐心指导！向老师致敬！

我真的差不多就要放弃问问题了，觉得自己好像没有必要追问，而且怕耽误后面同学问问题的时间。不过老师真的太强了，这种犄角旮旯的问题也能解决！老师的储备真的吓人。谢谢老师没有打击我，我问问题，舌头都捋不直，哎～师兄还说我在说啥，但是老师依旧认真在听我讲述，太感动了。

2021.11.11

Panel Test

Tang Jie, Zou Yunlong

2022 年 11 月 2 日

目录

1	Model	1
2	LR Test	2
3	Simulation	3
4	EL Test	6

1 Model

I_κ is an identity matrix of dimension κ , J_κ is a matrix of ones of dimension κ , $\mathbf{1}_\kappa$ is a vector of ones of dimension κ , \otimes denotes the Kronecker product.

Consider the following panel data regression model,

$$y_{it} = x'_{it}\beta + u_{it}, \quad i = 1, 2, \dots, N, \quad t = 1, 2, \dots, T, \quad (1)$$

where y_{it} is the observation on the i -th region for the t -th time period, x_{it} denotes the $k \times 1$ vector of observations on the non-stochastic regressors and u_{it} is the regression disturbance. The disturbance vector of (1) is assumed to have random region effects as well as spatially autocorrelated residual disturbances, i.e.,

$$u_t = \mu + \varepsilon_t \quad (2)$$

with

$$\varepsilon_t = \lambda W \varepsilon_t + v_t, \quad (3)$$

where $u'_t = (u_{1t}, u_{2t}, \dots, u_{Nt})$, $\varepsilon'_t = (\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{Nt})$ and $\mu' = (\mu_1, \mu_2, \dots, \mu_N)$ denote the vector of random region effects which are assumed to be $\text{IIN}(0, \sigma_\mu^2)$. λ satisfies $|\lambda| < 1$. W is known $N \times N$ spatial weight matrix. $v'_t = (v_{1t}, v_{2t}, \dots, v_{Nt})$, where v_{it} is i.i.d over i and t and is assumed to be $\text{N}(0, \sigma_v^2)$. One can rewrite (3) as

$$\varepsilon_t = (I_N - \lambda W)^{-1} v_t = B^{-1} v_t, \quad (4)$$

where $B = I_N - \lambda W$. Then (2) can be written as

$$u_t = \mu + B^{-1} v_t, \quad (5)$$

Denoting $y_t = (y_{1t}, y_{2t}, \dots, y_{Nt})'$, and $x_t = (x_{1t}, x_{2t}, \dots, x_{Nt})'$, the model (1) has the following reduced-form representation,

$$y_t = x_t \beta + u_t, \quad t = 1, 2, \dots, T, \quad (6)$$

We continue with the model (5)-(6). With $t = 1, 2, \dots, T$, model (5)-(6) can be written into a matrix form as follows:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_T \end{pmatrix} \beta + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{pmatrix}$$

with

$$\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{pmatrix} = \begin{pmatrix} \mu \\ \mu \\ \vdots \\ \mu \end{pmatrix} + \begin{pmatrix} B^{-1} & 0 & \cdots & 0 \\ 0 & B^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B^{-1} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_T \end{pmatrix}$$

or

$$y = X\beta + u, \quad (7)$$

with

$$u = (\mathbf{1}_T \otimes I_N)\mu + (I_T \otimes B^{-1})v, \quad (8)$$

where $y' = (y'_1, y'_2, \dots, y'_T)$, $X' = (x'_1, x'_2, \dots, x'_T)$, $u' = (u'_1, u'_2, \dots, u'_T)$, and $v' = (v'_1, v'_2, \dots, v'_T)$.

2 LR Test

To clearly show the model in this report, the model (7)-(8) is rewritten as follows

$$y = X\beta + u, \quad u = (\mathbf{1}_T \otimes I_N)\mu + (I_T \otimes B^{-1})v.$$

This variance-covariance matrix can be written as

$$\Omega_u = \sigma_v^2 \Sigma_u$$

where $\Sigma_u = \bar{J}_T \otimes [T\phi I_N + (B'B)^{-1}] + E_T \otimes (B'B)^{-1}$, with $\bar{J}_T = J_T/T$, $\phi = \sigma_\mu^2/\sigma_v^2$, $E_T = I_T - \bar{J}_T$. Σ_u^{-1} is given by

$$\Sigma_u^{-1} = \bar{J}_T \otimes [T\phi I_N + (B'B)^{-1}]^{-1} + E_T \otimes (B'B). \quad (9)$$

Under the assumption of normality, the log-likelihood function for this model was derived as

$$\begin{aligned} L &= -\frac{NT}{2} \log(2\pi\sigma_v^2) - \frac{1}{2} \log |\Sigma_u| - \frac{1}{2\sigma_v^2} u' \Sigma_u^{-1} u \\ &= -\frac{NT}{2} \log(2\pi\sigma_v^2) - \frac{1}{2} \log |T\phi I_N + (B'B)^{-1}| + \frac{(T-1)}{2} \log |B'B| - \frac{1}{2\sigma_v^2} u' \Sigma_u^{-1} u \end{aligned} \quad (10)$$

with $u = y - X\beta$.

Baltagi et al. (2003) derived the lagrange multiplier (LM) tests and the corresponding likelihood ratio (LR) tests for this model. Here, we redo the LR test under the hypotheses for $H_0^a : \lambda = \sigma_\mu^2 = 0$ and the alternative H_1^a is that at least one component is not zero.

Firstly, we obtain the estimations of the unrestricted log-likelihood function using the method of scoring. Let $\hat{\sigma}_v^2$, $\hat{\phi}$, $\hat{\lambda}$ and $\hat{\beta}$ denote the unrestricted maximum likelihood estimators (MLEs) and let $\hat{B} = I_N - \hat{\lambda}W$ and $\hat{u} = y - X'\hat{\beta}$, and then, by (10), the unrestricted maximum log-likelihood function is given by

$$L_U = -\frac{NT}{2} \log(2\pi\hat{\sigma}_v^2) - \frac{1}{2} \log|T\hat{\phi}I_N + (\hat{B}'\hat{B})^{-1}| + (T-1) \log|\hat{B}| - \frac{1}{2\hat{\sigma}_v^2} \hat{u}'\hat{\Sigma}_u^{-1}\hat{u},$$

where $\hat{\Sigma}_u^{-1}$ is obtained from (9) with $\hat{\beta}$ replacing β and $\hat{\phi}$ replacing ϕ .

Secondly, we obtain the estimations of the restricted log-likelihood function. Under the null hypothesis H_0^a , the restricted MLE of β is $\tilde{\beta}_{OLS}$, and $\tilde{\beta}_{OLS} = (X'X)^{-1}X'Y$, so that $\tilde{u} = y - X'\tilde{\beta}_{OLS}$ and $\tilde{\sigma}_v^2 = \tilde{u}'\tilde{u}/NT$. Therefore, the restricted maximum log-likelihood function under H_0^a is given by

$$L_R = -\frac{NT}{2} \log(2\pi\tilde{\sigma}_v^2) - \frac{1}{2\tilde{\sigma}_v^2} \tilde{u}'\tilde{u}.$$

Finally, the LR test statistic for H_0^a is given by

$$LR = 2(L_U - L_R) \quad (11)$$

and this should be asymptotically distributed as a mixture of χ^2 given in (12) under the null hypothesis H_0^a , χ_m^2 has a mixed χ^2 -distribution:

$$\chi_m^2 \sim \frac{1}{4}\chi^2(0) + \frac{1}{2}\chi^2(1) + \frac{1}{4}\chi^2(2) \quad (12)$$

The critical values for the mixed χ_m^2 are 7.289, 4.321 and 2.952 for $\alpha = 0.01, 0.05$ and 0.1 , respectively.

3 Simulation

We conducted a small simulation study to compare the finite sample performances of the critical regions based on LR tests at $\alpha = 0.05$ level, and report the proportion of $LR > \chi_m^2$ in our 2,000 simulations. The results of simulations are reported in Tables 1-2.

In the simulations, we used the model as Baltagi et al. (2003):

$$y_{it} = \beta_1 + x'_{it}\beta_2 + u_{it}, i = 1, \dots, N, t = 1, \dots, T,$$

where $\beta_1 = 5$, $\beta_2 = 0.5$ and $x_{it} = 0.1t + 0.5x_{i,t-1} + z_{it}$ with $z_{it} \sim U(-0.5, 0.5)$ and $x_{i0} = 5 + 10z_{i0}$. W is a queen type weight matrix. For the disturbances, $u_{it} = \mu_i + \varepsilon_{it}$, $\varepsilon_{it} = \lambda \sum_{j=1}^N w_{ij}\varepsilon_{jt} + v_{it}$

with $\mu_i \sim \text{IIN}(0, \sigma_\mu^2)$ and $v_{it} \sim \text{N}(0, \sigma_v^2)$. We fix $\sigma_\mu^2 + \sigma_v^2 = 20$ and let $\rho = \sigma_\mu^2 / (\sigma_\mu^2 + \sigma_v^2)$ vary over the set $(0, 0.2, 0.5)$. The spatial autocorrelation factor λ is varied over a positive range from 0 to 0.9 by increments of 0.1. Two values for $N = 25$ and 49, and two values for $T = 3$ and 7 are chosen.

表 1: LR test for H_0^a , $\lambda = \sigma_\mu^2 = 0$

N, T	λ	$\rho = 0$		$\rho = 0.2$		$\rho = 0.5$	
		Baltagi	Tang	Baltagi	Tang	Baltagi	Tang
25, 3	0.0	0.064	0.071	0.372	0.412	0.958	0.976
	0.1	0.066	0.100	0.380	0.421	0.967	0.969
	0.2	0.140	0.180	0.411	0.459	0.968	0.969
	0.3	0.282	0.316	0.519	0.564	0.981	0.977
	0.4	0.514	0.533	0.650	0.670	0.984	0.987
	0.5	0.743	0.740	0.806	0.845	0.992	0.989
	0.6	0.891	0.908	0.923	0.946	0.995	0.999
	0.7	0.974	0.977	0.980	0.985	1.000	0.999
	0.8	0.997	1.000	0.996	1.000	0.999	1.000
	0.9	1.000	1.000	1.000	1.000	1.000	1.000
25, 7	0.0	0.056	0.074	0.876	0.902	1	1.000
	0.1	0.117	0.107	0.904	0.910	1	1.000
	0.2	0.317	0.340	0.932	0.945	1	0.999
	0.3	0.646	0.670	0.951	0.973	1	1.000
	0.4	0.883	0.894	0.989	0.992	1	1.000
	0.5	0.987	0.989	0.999	0.999	1	1.000
	0.6	0.999	1.000	1.000	1.000	1	1.000
	0.7		1.000		1.000		1.000
	0.8		1.000		1.000		1.000
	0.9		1.000		1.000		1.000

Simulation results show that the frequency of rejections in Baltagi et al. (2003) is larger than Tang and Zou. Some blank means no data in Baltagi et al. (2003).

表 2: LR test for H_0^a , $\lambda = \sigma_\mu^2 = 0$

N, T	λ	$\rho = 0$		$\rho = 0.2$		$\rho = 0.5$	
		Baltagi	Tang	Baltagi	Tang	Baltagi	Tang
49, 3	0.0	0.057	0.035	0.622	0.599	1.000	0.998
	0.1	0.100	0.042	0.636	0.598	1.000	0.999
	0.2	0.232	0.075	0.716	0.635	1.000	1.000
	0.3	0.517	0.119	0.804	0.685	1.000	1.000
	0.4	0.782	0.235	0.901	0.762	1.000	1.000
	0.5	0.953	0.382	0.973	0.820	1.000	1.000
	0.6	0.993	0.600	0.995	0.894	1.000	0.999
	0.7	1.000	0.790	1.000	0.958	1.000	0.997
	0.8		0.907		0.982		0.985
	0.9		0.906		0.979		0.996
49, 7	0.0	0.056	0.025	0.995	0.992	1.000	1.000
	0.1	0.153	0.054	0.995	0.992	1.000	1.000
	0.2	0.536	0.198	0.996	0.996	1.000	1.000
	0.3	0.899	0.503	0.999	0.997	1.000	1.000
	0.4	0.994	0.777	1.000	0.998	1.000	1.000
	0.5	1.000	0.926	1.000	0.997	1.000	1.000
	0.6	1.000	0.992	1.000	0.998	1.000	1.000
	0.7		0.999		1.000		1.000
	0.8		1.000		1.000		1.000
	0.9		0.963		1.000		1.000

4 EL Test

ℓ is the likelihood function, L is the log-likelihood function (10):

$$LR = -2 \log \frac{\ell_R}{\ell_U} = 2(L_U - L_R) \quad (13)$$

$\tilde{\ell}$ is the empirical likelihood, $\tilde{\ell}(b) = \sup_{\{p_i\}_{i=1}^n} \prod_{i=1}^n p_i$, $\tilde{L}(b)$ is the empirical log-likelihood:

$$EL = -2 \log \frac{\sup_{b \in \Omega_0} \tilde{L}(b)}{\sup_{b \in \Omega_0 + \Omega_1} \tilde{L}(b)} = 2(\tilde{L}_{\Omega_0 + \Omega_1} - \tilde{L}_{\Omega_0}) \quad (14)$$

$$EL = -2 \log \frac{\sup_{b \in \Omega_0} \tilde{L}(b)}{\sup_{b \in \Omega_0 + \Omega_1} \tilde{L}(b)} = -2 \log \frac{\sup_{b \in \Omega_0} \tilde{L}(b)}{\sup_{b \in \mathbf{R}} \tilde{L}(b)} I(\hat{b} \in \Omega_1) \quad (15)$$

高维数据

唐洁, 邹云龙

2022 年 11 月 8 日

目录

1	空间模型	1
1.1	模型介绍	1
1.2	分析汇报	2
1.3	实验设计	2
2	模拟结果	3
2.1	实验一	3
2.2	实验二	5

1 空间模型

1.1 模型介绍

含空间自相关误差的空间自回归模型 (SARARmodel):

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

其中, n 是空间单元数量, $\rho_j, j = 1, 2$ 是空间自回归系数且 $|\rho_j| < 1, j=1, 2$, $X_n = (x_1, x_2, \dots, x_n)'$ 是 $n \times p$ 维解释变量的样本资料矩阵, β 是 $p \times 1$ 维 X_n 的回归系数向量, $Y_n = (y_1, y_2, \dots, y_n)'$ 是 $n \times 1$ 维响应变量, W_n 是解释变量 Y_n 的空间邻接权重矩阵, M_n 是扰动项 $u_{(n)}$ 的空间邻接权重矩阵, $\epsilon_{(n)}$ 是 $n \times 1$ 维空间误差向量, 且满足

$$E\epsilon_{(n)} = 0, \text{Var}(\epsilon_{(n)}) = \sigma^2 I_n.$$

本汇报将呈现 $p \rightarrow \infty$ 的情形. 符号说明, $\mathbf{1}_k$ 表示数字 1 组成的 k 维列向量.

由 Qin, 得到关于 $\theta = (\beta', \rho_1, \rho_2, \sigma^2)' \in R^{p+3}$ 的经验 (对数) 似然比统计量:

$$\ell_n(\theta) = -2 \log L_n(\theta) = 2 \sum_{i=1}^n \log \{1 + \lambda'(\theta) \omega_i(\theta)\},$$

$$L_n(\theta) = \sup \left\{ \prod_{i=1}^n (np_i) : p_1 \geq 0, \dots, p_n \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \omega_i(\theta) = 0 \right\},$$

其中

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

情形 1 当 p 固定时, 模型 (1) 下, 当 $\theta = \theta_0, n \rightarrow \infty$, 有

$$el = \ell_n(\theta_0) \xrightarrow{d} \chi_{p+3}^2.$$

情形 2 当 $p \rightarrow \infty$ 时, 模型 (1) 下, 当 $\theta = \theta_0, p = cn^{index}, n \rightarrow \infty$, 有

$$hel = \ell_n^h(\theta_0) \xrightarrow{d} \chi_1^2,$$

其中 $\ell_n^h(\theta) = -2 \log L_n^h(\theta) = 2 \sum_{i=1}^n \log \{1 + \lambda'(\theta) \omega_i^h(\theta)\}$, 且 $\omega_i^h(\theta) = \omega_i'(\theta) \mathbf{1}_{p+3}, i = 1, 2, \dots, n$.

1.2 分析汇报

在空间模型中, $p = cn^{index}$, 其中 p 为 X_n 的维数.

当 $index = 0$, 维数为固定常数, 由表 1-2 可看出, 情形 1 得到验证. 当 $index > 0$, $p \rightarrow \infty$, 由表 1-2 可看出, 情形 2 得到验证, 尤其是当 $index = 0.5$, 由表 1-2 可知, EL 的覆盖率都几乎为 0, 而 HEL 仍然接近名义水平. 换言之, HEL 犯第一类错误的频率比 EL 要低, 并且随着维数的增加, HEL 还很接近名义水平.

另外, 我们设计了功效函数实验, 比较二者犯第二类错误的表现. 表 3 可知, 随着参数与真值参数的距离越远, HEL 犯第二类错误的频率越接近于 0, 即功效函数越接近于 1. 而 EL 的表现劣于 HEL, 因为 EL 在参数与真值参数的距离较近的时候, 随着维数变大都趋于 1.

综上, HEL 比 EL 在第一类错误和第二类错误的表现中都好.

1.3 实验设计

我们通过模拟比较 EL 和 HEL 优劣, 进行模拟 1000 次, 在模拟中, 使用如下模型:

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

其中 $\rho_1 = 0.85$, $\rho_2 = 0.15$, $\beta_0 = \mathbf{1}_{p+3}$, $p = [cn^{index}]$, $[\cdot]$ 表示取整函数, $\{X_i\}$ 服从 $N(\mathbf{0}, \Sigma_P)$, 其中 $\Sigma_P = I_{\{i=j, 1 \leq i, j \leq p\}}$, ϵ_i s 分别来自 $N(0, \sigma_0^2)$, 其中 $\sigma_0^2 = 1$. 空间权重矩阵采用皇后邻接, 考虑空间单元的 4 种理想情况: 规则正方形网格 $n = m \times m$, $m=10, 15, 20, 30$ 分别表示 W_n 为 $grid_{100}$ 、 $grid_{225}$ 、 $grid_{400}$ 和 $grid_{900}$.

记 $\theta' = (\beta', \rho_1, \rho_2, \sigma^2)$, $\beta = \beta_0 + \Delta$, $\sigma^2 = \sigma_0^2 + \Delta$, Δ 取值依次为 0、0.1、1 以及 2.

原假设 $H_0: \theta = \theta_0$, 即 $\Delta = 0$. 备择假设 $H_1: \theta \neq \theta_0$, 即 $\Delta \neq 0$. 取定 α , $0 < \alpha < 1$, 设 $z_\alpha(p+3)$ 满足 $P(\chi_{p+3}^2 > z_\alpha(p+3)) = \alpha$, $z_\alpha(1)$ 满足 $P(\chi_1^2 > z_\alpha(1)) = \alpha$.

实验一: 原假设下置信区域覆盖率

$\Delta = 0$, $c = (3, 4, 5)$, $index = (0, 0.14, 0.16, 0.24, 0.3, 0.4, 0.5)$, 分别给出 $\ell_n(\theta_0) \leq z_\alpha(p+3)$ 和 $\ell_n^h(\theta_0) \leq z_\alpha(1)$ 在模拟中出现的比例, 其中 θ_0 是 θ 的真实值.

实验二: 备择假设下拒绝区域覆盖率

$\Delta = (0, 0.1, 1, 2)$, $c = 3$, $index = (0, 0.1, 0.2, 0.3, 0.4, 0.5)$, 分别给出 $\ell_n(\theta) \geq z_\alpha(p+3)$ 和 $\ell_n^h(\theta) \geq z_\alpha(1)$ 在模拟中出现的比例, 其中 $\theta \neq \theta_0$ 是 θ 的非真实值. 值得一提的是, 当 $\Delta = 0.1$ 时, $\|\theta - \theta_0\| \leq (n/p)^{-1/3}$.

2 模拟结果

2.1 实验一

表 1: $1 - \alpha = 0.90$

$p, \text{EL, HEL} \backslash n$ <i>index</i>	100	100	100	225	225	225	400	400	400	900	900	900
0	3	0.825	0.883	3	0.857	0.900	3	0.879	0.883	3	0.892	0.898
0	4	0.773	0.897	4	0.857	0.885	4	0.878	0.889	4	0.876	0.878
0	5	0.742	0.888	5	0.846	0.900	5	0.866	0.892	5	0.888	0.904
0.14	6	0.694	0.896	6	0.831	0.906	7	0.871	0.901	8	0.875	0.903
0.14	8	0.627	0.891	9	0.814	0.896	9	0.850	0.903	10	0.866	0.897
0.14	10	0.566	0.882	11	0.763	0.887	12	0.840	0.903	13	0.865	0.890
0.16	6	0.729	0.894	7	0.813	0.892	8	0.838	0.905	9	0.876	0.909
0.16	8	0.652	0.876	10	0.768	0.904	10	0.835	0.907	12	0.867	0.900
0.16	10	0.567	0.873	12	0.723	0.893	13	0.799	0.883	15	0.863	0.892
0.24	9	0.602	0.895	11	0.735	0.895	13	0.828	0.914	15	0.877	0.902
0.24	12	0.464	0.892	15	0.681	0.888	17	0.744	0.908	20	0.869	0.906
0.24	15	0.331	0.895	18	0.613	0.901	21	0.751	0.886	26	0.832	0.910
0.3	12	0.468	0.892	15	0.678	0.907	18	0.772	0.894	23	0.818	0.887
0.3	16	0.287	0.901	20	0.558	0.901	24	0.703	0.890	31	0.753	0.895
0.3	20	0.147	0.896	25	0.429	0.909	30	0.596	0.893	38	0.749	0.888
0.4	19	0.211	0.890	26	0.419	0.884	33	0.549	0.891	46	0.705	0.884
0.4	25	0.055	0.892	35	0.191	0.906	44	0.335	0.889	61	0.577	0.892
0.4	32	0.005	0.876	44	0.067	0.891	55	0.211	0.887	76	0.379	0.910
0.5	30	0.013	0.887	45	0.063	0.884	60	0.128	0.895	90	0.230	0.882
0.5	40	0.000	0.895	60	0.007	0.911	80	0.007	0.893	120	0.045	0.894
0.5	50	0.000	0.890	75	0.000	0.904	100	0.000	0.903	150	0.004	0.897

表 2: $1 - \alpha = 0.95$

$p, \text{EL, HEL} \backslash n$	100			225			400			900		
$index$	100	100	100	225	225	225	400	400	400	900	900	900
0	3	0.890	0.938	3	0.930	0.946	3	0.931	0.937	3	0.954	0.942
0	4	0.858	0.952	4	0.924	0.945	4	0.932	0.945	4	0.941	0.936
0	5	0.821	0.936	5	0.915	0.947	5	0.922	0.944	5	0.938	0.955
0.14	6	0.769	0.942	6	0.903	0.967	7	0.934	0.949	8	0.940	0.945
0.14	8	0.721	0.937	9	0.879	0.941	9	0.911	0.954	10	0.923	0.953
0.14	10	0.662	0.938	11	0.835	0.936	12	0.903	0.958	13	0.918	0.948
0.16	6	0.800	0.942	7	0.887	0.941	8	0.915	0.953	9	0.931	0.953
0.16	8	0.747	0.941	10	0.847	0.948	10	0.897	0.954	12	0.919	0.949
0.16	10	0.653	0.937	12	0.810	0.948	13	0.892	0.938	15	0.921	0.940
0.24	9	0.697	0.953	11	0.812	0.956	13	0.896	0.953	15	0.934	0.953
0.24	12	0.549	0.946	15	0.787	0.948	17	0.843	0.953	20	0.924	0.955
0.24	15	0.433	0.936	18	0.691	0.955	21	0.833	0.942	26	0.904	0.955
0.3	12	0.555	0.944	15	0.772	0.952	18	0.838	0.944	23	0.886	0.953
0.3	16	0.360	0.943	20	0.672	0.950	24	0.778	0.938	31	0.843	0.934
0.3	20	0.192	0.947	25	0.526	0.955	30	0.707	0.951	38	0.850	0.953
0.4	19	0.280	0.943	26	0.526	0.937	33	0.651	0.952	46	0.791	0.943
0.4	25	0.086	0.935	35	0.270	0.958	44	0.440	0.947	61	0.703	0.946
0.4	32	0.008	0.937	44	0.100	0.938	55	0.275	0.949	76	0.518	0.966
0.5	30	0.019	0.946	45	0.090	0.947	60	0.184	0.952	90	0.324	0.929
0.5	40	0.000	0.948	60	0.010	0.954	80	0.012	0.936	120	0.087	0.943
0.5	50	0.000	0.956	75	0.000	0.951	100	0.000	0.951	150	0.008	0.944

2.2 实验二

表 3: $1 - \alpha = 0.95$

N	p	$\Delta = 0$		$\Delta = 0.1$		$\Delta = 1$		$\Delta = 2$	
		EL	HEL	EL	HEL	EL	HEL	EL	HEL
100	3	0.146	0.058	0.205	0.091	0.994	0.083	1.000	0.976
	5	0.180	0.062	0.256	0.091	1.000	0.427	1.000	0.999
	8	0.282	0.066	0.325	0.082	1.000	0.732	1.000	1.000
	12	0.453	0.053	0.475	0.082	1.000	0.851	1.000	1.000
	19	0.755	0.058	0.726	0.065	1.000	0.891	1.000	1.000
	30	0.979	0.046	0.968	0.091	1.000	0.945	1.000	1.000
225	3	0.079	0.050	0.140	0.094	1.000	0.358	1.000	1.000
	5	0.097	0.044	0.129	0.089	1.000	0.907	1.000	1.000
	9	0.114	0.063	0.181	0.099	1.000	0.994	1.000	1.000
	15	0.220	0.060	0.263	0.100	1.000	1.000	1.000	1.000
	26	0.513	0.057	0.525	0.100	1.000	1.000	1.000	1.000
	45	0.925	0.045	0.911	0.108	1.000	1.000	1.000	1.000
400	3	0.072	0.061	0.146	0.141	1.000	0.735	1.000	1.000
	5	0.062	0.061	0.160	0.166	1.000	0.999	1.000	1.000
	10	0.092	0.047	0.143	0.117	1.000	1.000	1.000	1.000
	18	0.137	0.046	0.271	0.141	1.000	1.000	1.000	1.000
	33	0.342	0.048	0.586	0.142	1.000	1.000	1.000	1.000
	60	0.807	0.050	0.984	0.177	1.000	1.000	1.000	1.000
900	3	0.049	0.058	0.212	0.226	1.000	0.989	1.000	1.000
	6	0.050	0.055	0.168	0.221	1.000	1.000	1.000	1.000
	12	0.061	0.055	0.216	0.224	1.000	1.000	1.000	1.000
	23	0.101	0.046	0.552	0.235	1.000	1.000	1.000	1.000
	46	0.221	0.059	0.994	0.272	1.000	1.000	1.000	1.000
	90	0.654	0.049	1.000	0.334	1.000	1.000	1.000	1.000

高维数据

唐洁, 邹云龙

2022 年 11 月 15 日

目录

1	空间模型	1
1.1	模型介绍	1
1.2	新想法	1
2	主成分分析	2
2.1	原理	2
2.2	应用	3
3	PCAEL	4
3.1	主要结果	4
3.2	模拟	5

1 空间模型

1.1 模型介绍

含空间自相关误差的空间自回归模型 (SARARmodel):

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

其中, n 是空间单元数量, $\rho_j, j = 1, 2$ 是空间自回归系数且 $|\rho_j| < 1, j=1, 2$, $X_n = (x_1, x_2, \dots, x_n)'$ 是 $n \times p$ 维解释变量的样本资料矩阵, β 是 $p \times 1$ 维 X_n 的回归系数向量, $Y_n = (y_1, y_2, \dots, y_n)'$ 是 $n \times 1$ 维响应变量, W_n 是解释变量 Y_n 的空间邻接权重矩阵, M_n 是扰动项 $u_{(n)}$ 的空间邻接权重矩阵, $\epsilon_{(n)}$ 是 $n \times 1$ 维空间误差向量, 且满足

$$E\epsilon_{(n)} = 0, \text{Var}(\epsilon_{(n)}) = \sigma^2 I_n.$$

本汇报将呈现 $p \rightarrow \infty$ 的情形. 符号说明, $\mathbf{1}_k$ 表示数字 1 组成的 k 维列向量.

由 Qin, 得到关于 $\theta = (\beta', \rho_1, \rho_2, \sigma^2)' \in R^{p+3}$ 的经验 (对数) 似然比统计量:

$$\ell_n(\theta) = -2 \log L_n(\theta) = 2 \sum_{i=1}^n \log\{1 + \lambda'(\theta) \omega_i(\theta)\},$$

$$L_n(\theta) = \sup \left\{ \prod_{i=1}^n (np_i) : p_1 \geq 0, \dots, p_n \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i \omega_i(\theta) = 0 \right\},$$

其中

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

情形 1 当 p 固定时, 模型 (1) 下, 当 $\theta = \theta_0, n \rightarrow \infty$, 有

$$el = \ell_n(\theta_0) \xrightarrow{d} \chi_{p+3}^2.$$

情形 2 当 $p \rightarrow \infty$ 时, 模型 (1) 下, 当 $\theta = \theta_0, p = cn^{index}, n \rightarrow \infty$, 有

$$hel = \ell_n^h(\theta_0) \xrightarrow{d} \chi_1^2,$$

其中 $\ell_n^h(\theta) = -2 \log L_n^h(\theta) = 2 \sum_{i=1}^n \log\{1 + \lambda'(\theta) \omega_i^h(\theta)\}$, 且 $\omega_i^h(\theta) = \omega_i'(\theta) \mathbf{1}_{p+3}, i = 1, 2, \dots, n$.

1.2 新想法

秦老师提出, 既然 $\sum_{i=1}^n \omega_i(\theta)$ 可以由 $p+3$ 维通过 $\sum_{i=1}^n \omega_i'(\theta) \mathbf{1}_{p+3}$ 的方式将为 1 维, 那可不可以是 2 维, 3 维, $\dots, s(s \leq p+3)$ 维, 而且哏地降为 1 维也没有什么依据. 不如理解为是对得分函数做一个选择, 高维做选择, 自然就是主成分分析 (PCA) 了. 话不多说, 开干.

2 主成分分析

2.1 原理

a. 目标

我们假设将 p 维降为 s 维, 已知 $y \in R^p$, y (已标准化) 的协方差为 Σ . 进行主成分分析就是要找出 Z_p , 按照方差贡献度从多到少排序, 按照累计贡献率选出前 s 个.

$$y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{pmatrix}, \text{Cov}(y) = \Sigma \implies Z = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_p \end{pmatrix}, \text{Cov}(Z) = \tilde{\Sigma}.$$

b. 计算

首先由待定系数法可知, 求 Z 即求 A .

$$\begin{aligned} Z_1 &= a'_1 y = a_{11}Y_1 + a_{12}Y_2 + \cdots + a_{1p}Y_p \\ Z_2 &= a'_2 y = a_{21}Y_1 + a_{22}Y_2 + \cdots + a_{2p}Y_p \\ &\vdots \\ Z_p &= a'_p y = a_{p1}Y_1 + a_{p2}Y_2 + \cdots + a_{pp}Y_p \end{aligned} \implies \begin{aligned} Z &= Ay, \quad A' = (a_1, a_2, \cdots, a_p) \\ \tilde{\Sigma} &= \left(\text{cov}(Z_i, Z_j) \right)_{p \times p} = \left(a'_i \Sigma a_j \right)_{p \times p} \end{aligned}$$

其次对目标函数求解, 可得

$$\begin{aligned} Z_1 &= a'_1 y \quad a'_1 a_1 = 1 \quad \max a'_1 \Sigma a_1 & \max a'_1 \Sigma a_1 = \lambda_1 \quad a_1 = e_1 \\ Z_2 &= a'_2 y \quad a'_2 a_2 = 1 \quad \max a'_2 \Sigma a_2 & \max a'_2 \Sigma a_2 = \lambda_2 \quad a_2 = e_2 \\ &\vdots & \vdots \\ Z_p &= a'_p y \quad a'_p a_p = 1 \quad \min a'_p \Sigma a_p & \min a'_p \Sigma a_p = \lambda_p \quad a_p = e_p \end{aligned} \implies$$

其中, λ_i 是 Σ 的特征值且 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$, e_i 是对应特征值 λ_i 所对应的特征向量.

最后, 设 $V = \text{diag}(\lambda_1, \lambda_2, \cdots, \lambda_p)$, $E'_p = (e_1, e_2, \cdots, e_p)$, diag 表示由特征值组成的对角矩阵, 由上述我们可得:

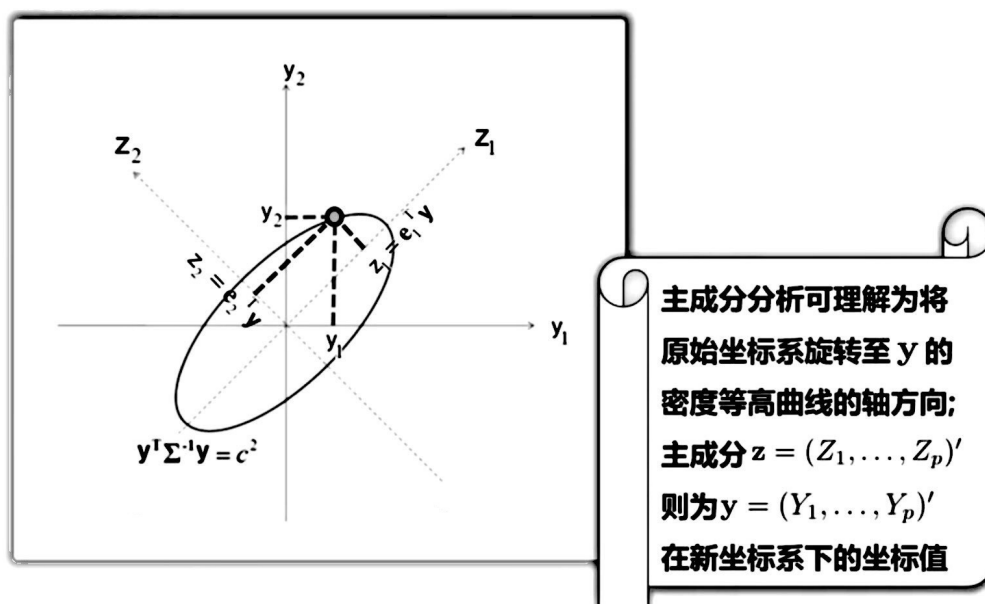
$$A = \begin{pmatrix} a'_1 \\ a'_2 \\ \vdots \\ a'_p \end{pmatrix} = \begin{pmatrix} e'_1 \\ e'_2 \\ \vdots \\ e'_p \end{pmatrix} = E_p$$

$$Z = Ay = E_p y$$

$$\tilde{\Sigma} = \left(a'_i \Sigma a_j \right)_{p \times p} = \left(e'_i \Sigma e_j \right)_{p \times p} = \left(\lambda_i I_{\{i=j\}} \right)_{p \times p} = V$$

c. 举例

例 如果 $y \sim N_2(0, \Sigma)$



d. 总结

按照 $Z = E_p y$ 的变换, 信息并没有损失, $\sum_{j=1}^p \text{Var}(Z_j) = \sum_{j=1}^p \lambda_j = \text{tr}(\Sigma) = \sum_{j=1}^p \text{Var}(Y_j)$.

第 i 个主成分的信息含量:

$$Z_i = e_i' y, \quad i = 1, 2, \dots, p.$$

第 i 个主成分的方差贡献度:

$$\lambda_i, \quad i = 1, 2, \dots, p.$$

第 i 个主成分的方差贡献度:

$$\frac{\lambda_i}{\sum_{j=1}^p \lambda_j}, \quad i = 1, 2, \dots, p.$$

前 s 个主成分的方差累积贡献度率:

$$\frac{\sum_{i=1}^s \lambda_i}{\sum_{j=1}^p \lambda_j}, \quad s \leq p.$$

假设我们从 p 维降到 s 维, 令 $E_s' = (e_1, e_2, \dots, e_s)$, 则做 $Z = E_s y$ 的变换, 此时 $Z \in R^s$. s 的选择可以根据方差累积贡献率进行判断.

2.2 应用

将上述结论应用到空间模型的得分函数选择上, 令 $y = \sum_{i=1}^n \omega_i(\theta)$, 则我们得出根据主成分分析得到得分函数选择结果为:

$$\omega_i^{PCA}(\theta) = E_s \omega_i(\theta), i = 1, 2, \dots, n,$$

其中 E_s 是 Σ_{p+3} 前 s 个最大特征值所对的特征向量组成的行特征向量矩阵.

3 PCAEL

3.1 主要结果

情形 **3** 当 $p \rightarrow \infty$ 时, 模型 (1) 下, 当 $\theta = \theta_0$, $n \rightarrow \infty$, 有

$$pcael = \ell_n^{PCAs}(\theta_0) \xrightarrow{d} \chi_s^2,$$

其中 $\ell_n^{PCAs}(\theta) = -2 \log L_n^{PCAs}(\theta) = 2 \sum_{i=1}^n \log\{1 + \lambda'(\theta) \omega_i^{PCAs}(\theta)\}$ 且 $\omega_i^{PCAs}(\theta) = E_s \omega_i(\theta)$, $i = 1, 2, \dots, n$,

$$E_s = \begin{pmatrix} e'_1 \\ e'_2 \\ \vdots \\ e'_s \end{pmatrix}, \quad \tilde{\Sigma}_s = Cov \left\{ \sum_{i=1}^n \omega_i^{PCA}(\theta) \right\} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_s \end{pmatrix}.$$

如此可将 $\omega_i(\theta)$ 从 $p+3$ 维降为 s 维, 其中 (λ_i, e_i) , $i = 1, 2, \dots, s$ 表示 Σ_{p+3} 的第 i 个特征值 λ_i 所对应的特征向量 e_i ,

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

其中,

$$\begin{aligned} \Sigma_{11} &= \sigma^2 \{B_n(\rho_2) X_n\}' B_n(\rho_2) X_n, \\ \Sigma_{12} &= \sigma^2 \{B_n(\rho_2) X_n\}' B_n(\rho_2) W_n A_n^{-1}(\rho_1) X_n \beta + \mu_3 \{B_n(\rho_2) X_n\}' Vec(diag \tilde{G}_n), \\ \Sigma_{13} &= \mu_3 \{B_n(\rho_2) X_n\}' Vec(diag \tilde{H}_n), \\ \Sigma_{14} &= \mu_3 \{B_n(\rho_2) X_n\}' \mathbf{1}_n, \\ \Sigma_{22} &= 2\sigma^4 tr(\tilde{G}_n^2) + \sigma^2 \{B_n(\rho_2) W_n A_n^{-1}(\rho_1) X_n \beta\}' B_n(\rho_2) W_n A_n^{-1}(\rho_1) X_n \beta \\ &\quad + (\mu_4 - 3\sigma^4) \|Vec(diag \tilde{G}_n)\|^2 + 2\mu_3 \{B_n(\rho_2) W_n A_n^{-1}(\rho_1) X_n \beta\}' Vec(diag \tilde{G}_n), \\ \Sigma_{23} &= 2\sigma^4 tr(\tilde{G}_n \tilde{H}_n) + (\mu_4 - 3\sigma^4) Vec'(diag \tilde{G}_n) Vec(diag \tilde{H}_n) \\ &\quad + \mu_3 \{B_n(\rho_2) W_n A_n^{-1}(\rho_1) X_n \beta\}' Vec(diag \tilde{H}_n), \\ \Sigma_{24} &= (\mu_4 - \sigma^4) tr(\tilde{G}_n) + \mu_3 \{B_n(\rho_2) W_n A_n^{-1}(\rho_1) X_n \beta\}' \mathbf{1}_n, \\ \Sigma_{33} &= 2\sigma^4 tr(\tilde{H}_n^2) (\mu_4 - 3\sigma^4) \|Vec(diag \tilde{H}_n)\|^2, \\ \Sigma_{34} &= (\mu_4 - \sigma^4) tr(\tilde{H}_n), \\ \Sigma_{44} &= n(\mu_4 - \sigma^4). \end{aligned}$$

3.2 模拟

取定 α , $0 < \alpha < 1$, 设 $z_\alpha(p+3)$ 满足 $P(\chi_{p+3}^2 > z_\alpha(p+3)) = \alpha$, $z_\alpha(1)$ 满足 $P(\chi_1^2 > z_\alpha(1)) = \alpha$ 以及 $z_\alpha(s)$ 满足 $P(\chi_s^2 > z_\alpha(s)) = \alpha$.

我们通过模拟比较 EL、HEL、和 PCAELs($s=1,2,3,4$) 优劣, 给定置信水平 $1 - \alpha = 0.95$, 分别给出 $\ell_n(\theta_0) \leq z_{0.05}(p+3)$ 、 $\ell_n^h(\theta_0) \leq z_{0.05}(1)$ 和 $\ell_n^{PCAs}(\theta_0) \leq z_{0.05}(s)$ 在 1000 次模拟中出现的比例, 其中 θ_0 是 θ 的真实值. 在模拟中, 使用如下模型:

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

其中 (ρ_1, ρ_2) 取为 $(0.85, 0.15)$, $\{X_i\}$ 服从 $N(\mathbf{0}, \Sigma_P)$, $\Sigma_P = I_{\{i=j, 1 \leq i, j \leq p\}}$, $\beta = \mathbf{1}_{p+3}$, $p = [cn^{index}]$, $[\cdot]$ 表示取整函数, $c = 4$, $index = 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7$ 和 0.8 , ϵ_i 's 分别来自 $N(0, 1)$, $N(0, 0.75)$, $t(5)$ 和 $\chi_4^2 - 4$, 其对应的三阶矩分别为 $\mu_3=0, 0, 0$ 和 32 , 四阶矩分别为 $\mu_4=3, \frac{9}{4}, 25$ 和 384 . 之所以选取这几个分布, 是因为这四个分布代表了标准正态分布、对称薄尾分布、对称厚尾分布以及非对称分布. 空间权重矩阵 $W_n = (w_{ij})$, w_{ij} 表示空间单元 i 与空间单元 j 之间的距离, 同一空间单元的距离 w_{ii} 为 0 , W_n 主对角元素为 0 , 显然, W_n 为对称矩阵. 模拟中, w_{ij} 的度量采用皇后邻接 (文献 [3] 第 18 页), 即:

$$w_{ij} = \begin{cases} 1, & \text{空间单元 } i \text{ 和 } j \text{ 有公共边或公共点,} \\ 0, & \text{其他.} \end{cases}$$

考虑空间单元的 4 种理想情况: 规则正方形网格 $n = m \times m$, $m=10, 15, 20$ 和 25 分别表示 W_n 为 $grid_{100}$ 、 $grid_{256}$ 、 $grid_{400}$ 和 $grid_{625}$.

表 1: EL、HEL 和 PCAELs ($s = 1, 2, 3, 4$) 置信域中的覆盖率于 $\epsilon_i \sim N(0, 1)$

$W_n = M_n$	$index$	p	EL	HEL	PCAEL1	PCAEL2	PCAEL3	PCAEL4
$grid_{100}$	0.0	4		0.927	0.932	0.935	0.915	0.898
	0.1	6		0.948	0.934	0.915	0.921	0.894
	0.2	10		0.941	0.936	0.929	0.920	0.905
	0.3	16		0.936	0.951	0.945	0.932	0.922
	0.4	25		0.938	0.948	0.948	0.930	0.913
	0.5	40		0.943	0.942	0.930	0.922	0.907
	0.6	63		0.950	0.953	0.927	0.921	0.909
	0.7	100		0.947	0.943	0.939	0.932	0.924
	0.8	159		0.933	0.925	0.925	0.917	0.896
	0.9	252		0.954	0.955	0.939	0.935	0.915
$grid_{225}$	0.0	4		0.942	0.950	0.937	0.927	0.936
	0.1	7		0.953	0.952	0.938	0.937	0.943
	0.2	12		0.953	0.954	0.946	0.933	0.937
	0.3	20		0.947	0.959	0.953	0.938	0.933
	0.4	35		0.942	0.944	0.937	0.936	0.937
	0.5	60		0.957	0.954	0.946	0.939	0.935
	0.6	103		0.951	0.952	0.946	0.949	0.944
	0.7	177		0.947	0.940	0.934	0.931	0.933
	0.8	305		0.944	0.945	0.949	0.951	0.940
	0.9	524		0.935	0.940	0.930	0.933	0.932
$grid_{400}$	0.0	4		0.945	0.941	0.940	0.933	0.935
	0.1	7		0.953	0.941	0.939	0.939	0.938
	0.2	13		0.952	0.952	0.940	0.938	0.937
	0.3	24		0.950	0.945	0.945	0.953	0.943
	0.4	44		0.953	0.951	0.939	0.943	0.939
	0.5	80		0.944	0.948	0.943	0.950	0.943
	0.6	146		0.943	0.956	0.954	0.953	0.951
	0.7	265		0.949	0.947	0.946	0.949	0.955
	0.8	483		0.945	0.955	0.941	0.943	0.939
	0.9	879		0.956	0.955	0.948	0.944	0.949
$grid_{665}$	0.0	4		0.953	0.955	0.952	0.947	0.939
	0.1	8		0.950	0.950	0.941	0.932	0.932
	0.2	14		0.946	0.950	0.955	0.951	0.951
	0.3	28		0.944	0.944	0.956	0.955	0.954
	0.4	53		0.943	0.948	0.949	0.941	0.940
	0.5	100		0.955	0.947	0.939	0.952	0.948
	0.6	190		0.952	0.961	0.951	0.953	0.955
	0.7	362		0.940	0.948	0.938	0.948	0.942
	0.8	690		0.953	0.955	0.944	0.941	0.938
	0.9	1313		0.953	0.956	0.942	0.946	0.941

表 2: EL、HEL 和 PCAELs ($s = 1, 2, 3, 4$) 置信域中的覆盖率于 $\epsilon_i \sim N(0, 0.75)$

$W_n = M_n$	$index$	p	EL	HEL	PCAEL1	PCAEL2	PCAEL3	PCAEL4
$grid_{100}$	0.0	4		0.942	0.941	0.930	0.922	0.914
	0.1	6		0.943	0.935	0.928	0.916	0.903
	0.2	10		0.953	0.945	0.933	0.929	0.919
	0.3	16		0.939	0.933	0.913	0.907	0.890
	0.4	25		0.952	0.942	0.932	0.914	0.909
	0.5	40		0.940	0.942	0.940	0.919	0.910
	0.6	63		0.942	0.939	0.926	0.923	0.915
	0.7	100		0.941	0.944	0.927	0.921	0.908
	0.8	159		0.943	0.930	0.931	0.919	0.906
	0.9	252		0.933	0.940	0.935	0.924	0.911
$grid_{225}$	0.0	4		0.941	0.944	0.943	0.929	0.919
	0.1	7		0.948	0.947	0.945	0.922	0.921
	0.2	12		0.944	0.946	0.949	0.944	0.934
	0.3	20		0.957	0.947	0.947	0.946	0.938
	0.4	35		0.937	0.939	0.933	0.939	0.920
	0.5	60		0.948	0.946	0.947	0.948	0.934
	0.6	103		0.942	0.948	0.948	0.947	0.939
	0.7	177		0.939	0.948	0.943	0.939	0.929
	0.8	305		0.944	0.949	0.942	0.941	0.935
	0.9	524		0.952	0.952	0.945	0.940	0.936
$grid_{400}$	0.0	4		0.955	0.939	0.944	0.941	0.940
	0.1	7		0.944	0.940	0.947	0.948	0.940
	0.2	13		0.957	0.957	0.956	0.960	0.956
	0.3	24		0.970	0.962	0.959	0.957	0.944
	0.4	44		0.936	0.948	0.947	0.943	0.945
	0.5	80		0.963	0.954	0.949	0.945	0.941
	0.6	146		0.947	0.943	0.950	0.940	0.947
	0.7	265		0.956	0.951	0.954	0.940	0.942
	0.8	483		0.941	0.939	0.936	0.933	0.940
	0.9	879		0.950	0.947	0.936	0.935	0.934

表 3: EL、HEL 和 PCAELs ($s = 1, 2, 3, 4$) 置信域中的覆盖率于 $\epsilon_i \sim t(5)$

$W_n = M_n$	$index$	p	EL	HEL	PCAEL1	PCAEL2	PCAEL3	PCAEL4
$grid_{100}$	0.0	4		0.916	0.904	0.887	0.867	0.837
	0.1	6		0.920	0.918	0.877	0.867	0.843
	0.2	10		0.923	0.909	0.874	0.847	0.840
	0.3	16		0.925	0.918	0.887	0.858	0.823
	0.4	25		0.914	0.924	0.890	0.869	0.853
	0.5	40		0.935	0.936	0.878	0.858	0.851
	0.6	63		0.949	0.934	0.873	0.852	0.830
	0.7	100		0.935	0.942	0.890	0.873	0.845
	0.8	159		0.924	0.935	0.879	0.862	0.837
	0.9	252		0.928	0.944	0.888	0.879	0.842
$grid_{225}$	0.0	4		0.932	0.920	0.914	0.911	0.898
	0.1	7		0.950	0.942	0.922	0.929	0.909
	0.2	12		0.936	0.948	0.923	0.915	0.906
	0.3	20		0.948	0.936	0.912	0.903	0.893
	0.4	35		0.932	0.950	0.918	0.911	0.903
	0.5	60		0.936	0.934	0.918	0.918	0.912
	0.6	103		0.950	0.948	0.924	0.925	0.903
	0.7	177		0.938	0.952	0.924	0.920	0.909
	0.8	305		0.948	0.919	0.907	0.888	0.892
	0.9	524		0.938	0.948	0.917	0.911	0.899
$grid_{400}$	0.0	4		0.944	0.923	0.935	0.923	0.925
	0.1	7		0.930	0.936	0.924	0.921	0.909
	0.2	13		0.941	0.930	0.920	0.911	0.904
	0.3	24		0.925	0.923	0.908	0.902	0.902
	0.4	44		0.946	0.937	0.918	0.906	0.905
	0.5	80		0.942	0.947	0.915	0.927	0.916
	0.6	146		0.943	0.949	0.914	0.912	0.898
	0.7	265		0.937	0.928	0.916	0.924	0.913
	0.8	483		0.949	0.954	0.918	0.912	0.901
	0.9	879		0.939	0.941	0.930	0.921	0.923
$grid_{625}$	0.0	4		0.930	0.926	0.936	0.935	0.930
	0.1	8		0.942	0.945	0.924	0.936	0.935
	0.2	14		0.944	0.937	0.933	0.937	0.935
	0.3	28		0.927	0.933	0.919	0.925	0.917
	0.4	53		0.943	0.945	0.945	0.939	0.931
	0.5	100		0.951	0.947	0.934	0.938	0.925
	0.6	190		0.953	0.950	0.920	0.917	0.917
	0.7	362		0.929	0.935	0.926	0.926	0.922
	0.8	690		0.945	0.931	0.935	0.921	0.918
	0.9	1313		0.948	0.944	0.938	0.934	0.940

表 4: EL 、HEL 和 PCAELs ($s = 1, 2, 3, 4$) 置信域中的覆盖率于 $\epsilon_i \sim \chi^2(4) - 4$

$W_n = M_n$	$index$	p	EL	HEL	PCAEL1	PCAEL2	PCAEL3	PCAEL4
$grid_{100}$	0.0	4		0.895	0.865	0.885	0.865	0.815
	0.1	6		0.915	0.890	0.885	0.870	0.855
	0.2	10		0.920	0.895	0.865	0.875	0.835
	0.3	16		0.920	0.920	0.915	0.880	0.875
	0.4	25		0.925	0.930	0.910	0.875	0.825
	0.5	40		0.955	0.950	0.925	0.900	0.860
	0.6	63		0.905	0.925	0.870	0.835	0.805
	0.7	100		0.935	0.935	0.900	0.860	0.870
	0.8	159		0.950	0.950	0.915	0.885	0.870
	0.9	252		0.940	0.935	0.875	0.840	0.855
$grid_{225}$	0.0	4		0.935	0.945	0.930	0.925	0.880
	0.1	7		0.910	0.955	0.955	0.905	0.880
	0.2	12		0.925	0.945	0.915	0.875	0.860
	0.3	20		0.950	0.920	0.925	0.900	0.885
	0.4	35		0.940	0.940	0.930	0.925	0.925
	0.5	60		0.955	0.945	0.930	0.900	0.890
	0.6	103		0.905	0.920	0.900	0.885	0.880
	0.7	177		0.935	0.920	0.935	0.895	0.890
	0.8	305		0.955	0.955	0.925	0.900	0.905
	0.9	524		0.955	0.920	0.945	0.900	0.895