

DOI: 10.13718/j.cnki.xdzk.2017.02.017

基于稳健主成分估计的建城区 面积变化驱动力分析

——以重庆市为例^①

张亚利^{1,2}, 唐菲菲^{1,2}

1. 重庆大学 土木工程学院, 重庆 400045; 2. 重庆大学 山地城镇建设与新技术教育部重点实验室, 重庆 400045

摘要: 建成区面积变化的驱动力分析是近年来土地利用领域的研究热点. 针对动力系统多重共线性及样本粗差问题, 将主成分估计和稳健估计有机结合, 构成稳健主成分估计, 并以重庆市建成区面积变化的驱动分析为例, 对稳健主成分估计方法的有效性进行了验证与分析. 研究表明: 当样本含有粗差时, 稳健主成分方法得出的驱动规律更科学、客观, 说明稳健主成分方法系统性更强, 对粗差有一定的抑制作用.

关键词: 主成分; 驱动力; 稳健估计; 重庆市

中图分类号: F301

文献标志码: A

文章编号: 1673-9868(2017)02-0107-07

20 世纪 90 年代以来, 土地利用变化的动力机制是目前国际上全球变化研究的前沿和热点课题之一, 学者们为此探索出了一些非常有效的研究方法^[1-6]. 其中, 基于经验的统计模型方法, 不仅有助于从复杂的土地利用系统中分离出主要的驱动因子, 确定土地利用变化与驱动因子的定量关系, 而且建模过程较简单, 限制条件少, 因此在土地利用驱动机制研究中广泛应用. 建成区面积变化驱动力分析就是土地利用驱动机制研究中的重要方面.

主成分分析在国内的土地利用变化驱动机制研究中应用较广^[5-8], 但因未构建定量模型, 不是完整意义上的定量研究. 逐步回归虽然属于定量研究, 但因逐步回归存在自变量的取舍, 破坏了驱动力系统的整体性, 不利于系统分析, 且未考虑样本(因变量)自身可能含有的粗差对参数解算的影响.

目前, 我国各地对建成区面积统计在统计方法、统计口径上并未统一, 即使同一地区, 二者也会有变化. 如 2012 年, 重庆市建委对用地面积分类做了调整. 因统计数据来源复杂, 产生漏算和重复计算的可能性较大, 且受统计人员的人为影响大, 容易产生粗差.

本文以重庆市建成区面积变化的驱动机制分析为例, 运用稳健主成分估计方法求解回归模型参数. 不仅可以利用主成分消弱驱动力系统多重共线性的影响, 而且可以有效抑制样本粗差对参数估计的影响, 使参数估计更加稳健, 回归模型更加科学、可靠, 符合现代地理学研究通过不断引入现代技术和数学模型进行深层次定量研究的发展趋势.

1 稳健主成分估计原理

Gauss-Markov 模型见公式(1):

$$\underset{n \times 1}{y} = \underset{n \times p}{X} \underset{p \times 1}{\beta} + \underset{n \times 1}{\varepsilon} \quad (1)$$

① 收稿日期: 2015-12-03

基金项目: 国家“十二五”科技支撑计划(2012BAJ22B00); 国家自然科学基金项目(41401380).

作者简介: 张亚利(1971-), 女, 内蒙古赤峰人, 副教授, 博士, 主要从事资源环境与数据处理方面的研究.

其中: $\varepsilon \sim (0, \sigma^2)$.

$X'X$ 所对应的特征根为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, (X' 为矩阵 X 的转置矩阵) 各特征根所对应的标准化的正交特征向量为 $a_1, a_2, \dots, a_p$. 则式(1) 的典则形式为:

$$y = F\alpha + \varepsilon \quad (2)$$

其中:

$$\begin{aligned} F &= XA & \alpha &= A'\beta \\ A &= (a_1, a_2, \dots, a_p) & F &= (F_1, F_2, \dots, F_p) \end{aligned}$$

记

$$\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_p)$$

令

$$\lambda_{m+1} \approx \lambda_{m+2} \approx \dots \approx \lambda_p \approx 0$$

为计算方便, 将 Λ, α, F, A 进行矩阵分块, 得

$$\begin{aligned} \Lambda &= \begin{pmatrix} \Lambda_1 & & \\ & \Lambda_2 & \\ & & \end{pmatrix} & \alpha &= \begin{pmatrix} \alpha_{(1)} \\ \alpha_{(2)} \end{pmatrix} \\ & \begin{matrix} m \times m \\ (p-m) \times (p-m) \end{matrix} & & \begin{matrix} m \times 1 \\ (p-m) \times 1 \end{matrix} \\ F &= \begin{pmatrix} F_{(1)} & F_{(2)} \\ n \times m & n \times (p-m) \end{pmatrix} & A &= \begin{pmatrix} A_{(1)} & A_{(2)} \\ p \times m & p \times (p-m) \end{pmatrix} \end{aligned}$$

因此, 公式(1) 的典则形式又变为:

$$y = F_{(1)}\alpha_{(1)} + F_{(2)}\alpha_{(2)} + \varepsilon \quad (3)$$

只保留公式(3) 中包含信息量最大的主成分项 $F_{(1)}\alpha_{(1)}$, 求得 $\alpha_{(1)}$ 的最小二乘估计

$$\hat{\alpha}_{(1)} = \Lambda_1^{-1} F_{(1)}' y \quad (4)$$

而

$$\alpha = A'\beta \quad (5)$$

则回归参数 β 的主成分估计

$$\tilde{\beta} = A \begin{pmatrix} \hat{\alpha}_{(1)} \\ 0 \end{pmatrix} = (A_{(1)} \quad A_{(2)}) \begin{pmatrix} \hat{\alpha}_{(1)} \\ 0 \end{pmatrix} = A_{(1)} \Lambda_1^{-1} A_{(1)}' X' y \quad (6)$$

稳健主成分估计不仅顾及系数矩阵 X 的多重共线性, 还要顾及 y 含有的粗差对 β 估计的影响.

根据 M 估计原理, 公式(1) 中 β 的稳健估计为:

$$\hat{\beta}_R = (X' \bar{P} X)^{-1} X' \bar{P} y \quad (7)$$

其中: $\bar{P}$ 为等价权阵.

令

$$\bar{N} = X' \bar{P} X$$

其特征根 $\bar{\lambda}_1 \geq \bar{\lambda}_2 \geq \dots \geq \bar{\lambda}_p$, 对应的标准正交化特征向量 $\bar{A} = (\bar{a}_1, \bar{a}_2, \dots, \bar{a}_p)$.

因 X 呈现多重共线性, 则 $\bar{N}$ 的特征根中 $\bar{\lambda}_{m+1} \approx \bar{\lambda}_{m+2} \approx \dots \approx \bar{\lambda}_p \approx 0$, 这表明, 它们所对应的主成分对模型影响很小, 可以剔除.

类似主成分估计的推导, 将 $\bar{A}, \bar{\Lambda}$ 分块得:

$$\bar{A} = \begin{pmatrix} \bar{\Lambda}_1 & & \\ & \bar{\Lambda}_2 & \\ & & \end{pmatrix} & \bar{A} &= \begin{pmatrix} \bar{A}_{(1)} & \bar{A}_{(2)} \\ p \times m & p \times (p-m) \end{pmatrix}$$

其中:

$$\bar{\Lambda} = \text{diag}(\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_p)$$

公式(1)中 β 的稳健主成分估计为:

$$\tilde{\beta}_R = \overline{A_1} \overline{A_1}^{-1} \overline{A_1}' X' \overline{P_y} \quad (8)$$

2 实例验证

2.1 研究区选择与驱动变量优选

重庆市自1997年直辖以来,发展迅速,特别是进入新世纪以来,随着国家“西部大开发”步伐的加快,重庆市的城乡面貌和社会经济得到了加速发展.重庆市地处三峡库区,地质环境破碎,生态环境脆弱,具有“大城市、大农村”的特点.近年来,重庆市不仅是土地利用变化研究的热点地区,也是城镇化发展研究的热点地区,因此本文选择重庆市作为研究区.

驱动建城区面积变化的动因很多,不仅有社会的、自然的因素,还有经济的、人文的等各方面的因素.土地利用变化驱动机制模型研究中,根据地球系统科学理论优选对其有重要影响的驱动力因子作为自变量,构成动力系统.自变量太多,不仅容易导致过度参数化,而且因为自变量之间的耦合关联,使分析者难以抓住关键要素,从而降低了驱动机制分析的针对性和有效性.因而,为减少驱动力之间的“相互驱动”和“层层驱动”,驱动力指标选取本着有代表性、权威、可解释、独立的原则,参考其他研究成果^[7, 9-12],优选了总户数 X_1 (万户)对应人口因素;固定资产投资额 X_3 (亿元)、国内生产总值 X_4 (亿元)、城镇居民人均可支配收入 X_5 (元)对应经济发展因素.为研究政策因素,尤其是户籍政策对建成区面积变化的推动作用,优选了统计方法、统计口径保持较为一致,数据权威性较高的户籍非农人口 X_2 (万人),与上述4个指标共同构成动力系统.建成区面积 Y (平方公里)作为因变量(样本).所有数据均来自2002年—2012年《重庆市统计年鉴》.

2.2 回归参数的稳健主成分方法求解

根据上述分析,建立影响重庆市建成区面积变化的回归模型:

$$Y = \beta_0 + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + X_4\beta_4 + X_5\beta_5 \quad (9)$$

为削弱变量数据数量级差异太大对模型求解的影响,提高驱动机制分析的可靠性,分别将变量 Y_i 和 $X = (X_{ij})_{11 \times 5}$ 标准化:

$$y_i = \frac{Y_i - Y^*}{S_Y^*} \quad x_{ij} = \frac{X_{ij} - X_j^*}{S_{X_j}^*} \quad i = 1, 2, \dots, 11; j = 1, 2, \dots, 5 \quad (10)$$

其中:

$$Y^* = \frac{\sum_{i=1}^{11} Y_i}{11} \quad S_Y^* = \sqrt{\frac{\sum_{i=1}^{11} (Y_i - Y^*)^2}{11 - 1}}$$
$$X_j^* = \frac{\sum_{i=1}^{11} X_{ij}}{11} \quad S_{X_j}^* = \sqrt{\frac{\sum_{i=1}^{11} (X_{ij} - X_j^*)^2}{11 - 1}}$$

用标准化变量表示的回归模型为:

$$y_i = x_{i1}\beta_1^* + x_{i2}\beta_2^* + x_{i3}\beta_3^* + x_{i4}\beta_4^* + x_{i5}\beta_5^* \quad (11)$$

以下对参数的求解,就是指对公式(11)中 $\beta_1^*, \beta_2^*, \beta_3^*, \beta_4^*, \beta_5^*$ 这5个参数的求解.

对于式(11),

$$\Lambda = \text{diag}(49.216\ 9, 0.530\ 4, 0.214\ 7, 0.029\ 7, 0.008\ 3)$$

条件数

$$k = \sqrt{\frac{\lambda_{\max}}{\lambda_i}} = \sqrt{\frac{49.216\ 9}{0.008\ 3}} = 77$$

按照累积贡献率

$$\delta = \frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^p \lambda_i} \geq 99\%$$

确定所选主成分个数为 2.

稳健主成分估计等价权采用下列函数:

$$\overline{P} = \begin{cases} P & |u| < k_0 \\ P \frac{k_0}{u} & k_0 \leq |u| \leq k_1 \\ 0 & |u| > k_1 \end{cases} \tag{12}$$

式中: P 为先验权; u 为标准化残差; $k_0 = 1.5$, $k_1 = 2.5$.

不加粗差时, 分别应用普通最小二乘方法、主成分方法和稳健主成分方法求解式(11)的回归参数, 数值如表 1 所示.

表 1 因变量不含粗差时的回归参数

方法	回 归 参 数				
普通最小二乘	2.187 0	0.083 2	-0.638 5	-0.506 5	-0.139 4
主成分	0.226 2	0.061 3	0.194 2	0.179 8	0.336 0
稳健主成分	0.220 0	0.075 5	0.195 5	0.181 5	0.326 7

回归参数具有“权重”意义: 参数值为正, 表明该参数所对应的自变量与因变量同向变化, 即“正相关”; 参数为负, 表明该参数所对应的自变量与因变量反向变化, 即“负相关”. 参数绝对值的大小表示自变量对因变量的作用强度: 绝对值越大, 表示该参数所对应的自变量对因变量的影响作用相对越强.

由表 1 可知, ① 在因变量不含有粗差时, 运用普通最小二乘方法求解回归参数, 对应于固定资产投资额、国内生产总值和城镇居民人均可支配收入的回归参数都出现了负值, 这与三者随建成区面积同向扩张的客观事实明显不符, 严重违背实际变化规律. 因此, 当系统存在较强的多重共线性时, 应用普通最小二乘方法求解回归参数不可行. ② 在因变量不含有粗差时, 运用主成分方法和稳健主成分方法所求各回归参数均为正值, 表明各驱动因子与建成区面积同向变化, 符合客观事实. 2 种方法所求参数数值基本一致, 最大相对误差为 $\frac{|0.336\,0 - 0.326\,7|}{0.336\,0} = 0.028$, 参数的 2 范数仅相差 0.003. 因此在样本没有粗差时, 2 个模型的精度没有区别.

为验证稳健主成分方法的抗差性, 在原始数据因变量 Y 中加入粗差, 即对 2002 年因变量加 900 km^2 (即 $\Delta_2 = 900$), 2007 年因变量加 900 km^2 (即 $\Delta_7 = 900$), 2002 年和 2007 年因变量同时加 900 km^2 (即 $\Delta_2 = \Delta_7 = 900$), 分别用主成分方法和稳健主成分方法求解回归参数. 为观察求解参数的稳健性, 计算它们与表 1 中的主成分方法所求参数的差的平方和的开方 D , D 值越小, 说明参数整体越稳健. 数据计算结果如表 2 所示.

表 2 因变量含粗差时的回归参数和 D

	主成分方法			稳健主成分方法		
	$\Delta_2 = 900$	$\Delta_7 = 900$	$\Delta_2 = \Delta_7 = 900$	$\Delta_2 = 900$	$\Delta_7 = 900$	$\Delta_2 = \Delta_7 = 900$
回归参数	0.029 7	0.301 2	0.166 4	0.174 7	0.160 0	0.150 5
	0.550 4	-0.744 0	-0.340 2	0.108 2	-0.018 7	-0.206 1
	0.128 3	0.102 2	0.069 9	0.165 2	0.150 2	0.104 9
	0.178 9	0.003 1	0.022 0	0.158 3	0.120 7	0.059 7
	-0.334 9	1.024 5	0.516 5	0.214 7	0.262 6	0.380 1
D	0.855 7	1.080 7	0.487 6	0.144 6	0.147 0	0.318 7

表 2 显示:

1) 当 $\Delta_2=900$ 时:

主成分方法求解 X_5^* 的回归参数为 $-0.334\ 9$, 表明人均可支配收入与建成区面积反向变化, 与二者同向变化的客观事实不符, 且参数的相对误差为 $\frac{|0.336\ 0+0.334\ 9|}{0.336\ 0}=1.997$; 而稳健主成分方法求解 X_5^* 的回归参数为 $0.214\ 7$, 相对误差为 $\frac{|0.336\ 0-0.214\ 7|}{0.336\ 0}=0.361\ 0$, 二者精度相差 5.53 倍.

稳健主成分方法所求各回归参数均为正值, 参数从大到小的排列次序为 $\tilde{\beta}_{R5}^*, \tilde{\beta}_{R1}^*, \tilde{\beta}_{R3}^*, \tilde{\beta}_{R4}^*, \tilde{\beta}_{R2}^*$, 与表 1 中不含粗差时主成分方法求解的参数大小排列规律完全相同.

主成分方法 $D=0.855\ 7$, 是稳健主成分方法 $D=0.144\ 6$ 的 5.92 倍.

2) 当 $\Delta_7=900$ 时:

主成分方法中, $\tilde{\beta}_2^*=-0.744\ 0$, 表明非农人口数与建成区面积反向变化, 相对误差为 $\frac{|-0.744\ 0-0.061\ 3|}{0.061\ 3}=13.137$. 而稳健主成分方法中 $\tilde{\beta}_{R2}^*=-0.018\ 7$, 相对误差为 $\frac{|-0.018\ 7-0.061\ 3|}{0.061\ 3}=1.30\ 5$, 可见二者的相对精度相差 10 倍之多; 主成分方法 $D=1.080\ 7$, 是稳健主成分方法 $D=0.147\ 0$ 的 7.35 倍.

3) 当 $\Delta_2=\Delta_7=900$ 时:

主成分方法中, $\tilde{\beta}_2^*=-0.340\ 2$, 相对误差为 $\frac{|-0.340\ 2-0.061\ 3|}{0.061\ 3}=6.549\ 8$; 而稳健主成分方法中, $\tilde{\beta}_{R2}^*=-0.206\ 1$, 其相对误差为 $\frac{|-0.206\ 1-0.061\ 3|}{0.061\ 3}=4.362\ 2$, 二者的相对精度相差 1.5 倍; 主成分方法 $D=0.487\ 6$, 是稳健主成分方法 $D=0.318\ 7$ 的 1.53 倍.

4) 等价权是决定稳健估计效果的关键, 由于 2002 年和 2007 年样本所含的多余观测分量不同, 所以当两者加粗差时, 稳健估计的效果会有所不同. 样本所含粗差个数越少, 稳健主成分方法的优越性越明显.

综合比较说明, 当样本含有粗差时, 稳健主成分方法与主成分方法相比较, 稳健主成分方法所求参数稳健性更强, 受样本粗差影响相对较小, 因此所建模型解释性和整体性更强, 结论更可靠.

2.3 驱动力分析

2.3.1 人口因素

土地利用是一个开放的系统, 人是最活跃的因素, 不仅是土地利用系统的组织者和参与者, 也是系统输出产品的消费者. 因此, 人口规模是城镇土地利用规模的决定性因素. 人口增长, 客观上要求对工业、商业、交通、住房等生产、生活用地增加供应, 推动了土地城市化进程. 人口增长越快, 城市面积的扩充也越快. 2001 年—2011 年, 重庆市总户数由 950.96 万户迅速增长到 1 205.20 万户, 增长了 26.8%. 人口的增长是建成区面积快速增加的重要驱动力.

2.3.2 经济因素

社会经济的发展是土地利用结构及其演变的主要动力. 重庆市自 2001 年至 2011 年 11 年间, 社会经济得到了长足发展. 在优选的 5 个驱动因子中, 城镇居民人均可支配收入是城区面积增长的最强驱动力, 人均可支配收入由 2001 年的 6 572 元迅猛增加到 2011 年的 20 250 元, 增长了 208%. 收入的增加, 使得人们对生活品质的改善需求旺盛, 居民住房、文化休闲、公共事业等需求用地的增加, 客观上要求城市扩展空间. 此外, 房产作为一种非常具有保值增值潜力的投资品, 人均可支配收入的增加, 意味着个人财富的增长, 不仅促进了自住房和生产生活环境的改善, 而且带动了房产投资, 促使建成区面积不断扩大.

在 2001 年至 2011 年的 11 年间,重庆市全社会固定资产投资额从 801.82 亿元增加到 7 685.87 亿元,增幅高达 858.55%;国内生产总值从 1 976.86 亿元增加至 10 011.13 亿元,增幅达 406.41%。固定资产投资额和国内生产总值的大幅增长,促进了“大城市带动大农村”的城镇化发展战略实施,导致基础设施建设和房地产业迅速崛起,建成区面积迅速增加。

2.3.3 政策因素

任何区域土地利用状态都是在特定的经济系统和政策水平下形成的。在所选的 5 个指标构成的动力系统中,总户数呈较强推动,而非农人口则相对弱些。一方面,非农人口的增长受户籍政策、农业农村政策影响较大。近年来,随着国家重视“三农”,有些农业人口会享受到一些补贴和收益,如三峡库区补贴,宅基地及农房、承包地的收益等,使一些人主观上不愿意放弃农村户口。另一方面,我国的户籍政策改革步伐还落后于经济发展,农业人口落户城市还存在体制机制上的诸多障碍。因此,在与总户数、固定资产投资额、国内生产总值和城镇居民人均可支配收入的对比中,非农人口的动力作用处于相对弱势。

3 结论与讨论

按照系统科学理论,优选总户数、固定资产投资额、国内生产总值、城镇居民人均可支配收入和非农人口数构成动力系统,运用稳健主成分分析方法,建立重庆市 2001 年—2011 年建成区面积变化的驱动模型。当建成区面积含有粗差时,稳健主成分方法得出的驱动规律较主成分方法得出的驱动规律更科学、客观,说明稳健主成分方法对粗差有一定的抑制作用,是一种非常有效的系统分析方法。当建成区面积不含有粗差时,稳健主成分与主成分构建的模型精度无差异。

2001 年—2011 年,重庆市建成区面积大幅增加。在所构成的动力系统中,驱动作用由强到弱依次为城镇居民人均可支配收入、总户数、固定资产投资额、国内生产总值、非农人口数。对于由其他指标构成的动力系统,其驱动强弱对比可能会有所不同。

可能实际观测中因变量的某些峰值并不是粗差,而正好体现的是相关变量间的某种不具有普遍代表意义的特殊规律。考虑到回归分析的目的是找出相关变量间的普遍规律,因而也将这种峰值用稳健估计方法进行处理。稳健主成分方法的抗差效果主要取决于等价权,对于不同类型的样本数据,采用哪种方式确定等价权将是下一步的研究重点。

参考文献:

- [1] 杨 梅,张广录,侯永平.区域土地利用变化驱动力研究进展与展望[J].地理与地理信息科学,2011,27(1):95—100.
- [2] SCHILLING K E, WOLTER C F. A GIS-Based Groundwater Travel Time Model to Evaluate Stream Nitrate Concentration Reductions From Land Use Change[J]. Environ Geol, 2007, 53: 433—443.
- [3] LI Zheng-hai, SONG Huo-bao, LU Hai-yan, et al. Analysis of Land Use Change and Its Driving Force in The Longitudinal Rang-Gorge Region[J]. Chinese Science Bulltine, 2007, 52: 10—20.
- [4] QUAN Bin, M J M RÖMKENS, TAO Jian-jun, et al. Spatial-Temporal Pattern and Population Driving Force of Land Use Change in Liupan Mountains Region, Southern Ningxia, China[J]. Chin Geogra Sci, 2008 18(4): 323—330.
- [5] 张小虎,张合兵,雷国平.大庆市土地利用变化及驱动机制分析[J].干旱区资源与环境,2012,26(9):109—113.
- [6] 高啸峰,王树德,宫阿都,等.基于主成分分析法的土地利用/覆被变化驱动力研究[J].地理与地理信息科学,2009,25(1):36—39.
- [7] 王小玉,张安明,邹小红,等.重庆市土地利用变化及驱动机制研究[J].西南大学学报(自然科学版),2009,31(2):146—151.
- [8] 马建辉,刘淑苹.基于 SPSS 的莆田市土地利用变化及其驱动力分析[J].长春大学学报,2009,19(10):71—74.
- [9] 范科红,李阳冰.1986—2007 年重庆主城区土地利用变化及其驱动因素研究[J].水土保持研究,2012,19(1):168—173.

[10] 杨 明, 郭仁忠, 李 全. 快速城市化地区建设用地演变驱动力研究——以深圳市龙岗区为例 [J]. 武汉大学学报(信息科学版), 2009, 34(2): 54—59.

[11] 陈红顺, 夏 斌. 快速城市化地区土地利用变化及驱动因素分析——以广东省东莞市为例 [J]. 水土保持通报, 2012, 32(1): 54—59.

[12] 毛蒋兴, 闫小培, 李志刚, 等. 快速城市化过程中深圳土地利用变化的自然及人文因素综合研究 [J]. 自然资源学报, 2009, 24(3): 523—535.

Driving Force Analysis of Urban Built-up Area Change Based on Robust Principal Component Estimation ——Taking Chongqing City as an Example

ZHANG Ya-li^{1,2}, TANG Fei-fei^{1,2}

- 1. School of Civil Engineering, Chongqing University, Chongqing 400045, China;
- 2. Key Laboratory of New Technology for Construction of Cities in Mountain Area of Ministry of Education, Chongqing University, Chongqing 400045, China

Abstract: Driving force analysis of urban built-up area change has become a primary type of land use change. According to multicollinearity of driving forces and block errors of samples, robust principal component estimation is established based on principal component estimation and robust estimation. Taking urban built-up area change of Chongqing city as an example, validity of robust principal component estimation is checked and analyzed. The results show that, as some block errors are in samples, the rule derived from robust principal component estimation is more scientific and objective, so robust principal component estimation is more systematic and brings down the influence of block errors.

Key words: principal component; driving force; robust estimation; Chongqing city

责任编辑 胡 杨

