

DOI:10.13718/j.cnki.xsxb.2023.03.010

基于 SAIC 方法的纵向数据模型平均^①

王梓屹

西南大学 数学与统计学院, 重庆 400715

摘要: 传统的 SAIC 模型平均所需运行时间随数据维数而呈现出阶乘级的增长, 其预测精度也随之下降. 本文基于传统 SAIC 模型平均法进行了改进, 提出一类基于 SAIC 加权法的纵向数据模型平均法, 使运算效率大幅提升, 并且使预测效果拥有良好的稳定性. 模拟实验结果表明, 与传统方法相比, 在预测残差平方和层面, 本文提出的新模型在稳定性、精准性和运行速度方面均优于传统方法.

关键词: 大数据; 赤池信息量准则; 模型平均; 广义估计方程; S-AIC 模型平均

中图分类号: O212.4

文献标志码: A

文章编号: 1000-5471(2023)03-0066-08

Longitudinal Data Model Averaging Based on SAIC Method

Wang Ziyi

School of Mathematics and Statistics, Southwest University, Chongqing 400715, China

Abstract: The running time required for traditional SAIC model averaging shows a stepwise increase with the number of data dimensions, and its prediction accuracy decreases. In this paper, we propose a class of longitudinal data model averaging method based on the traditional SAIC model averaging method, which can significantly improve the computational efficiency and make the prediction results have good stability. The simulation experimental results show that the new model proposed in this paper outperforms the traditional method in terms of stability, accuracy and running speed in predicting the sum of squared residuals compared with the traditional method.

Key words: big data; Akaike's information criterion; model averaging; generalized estimating equation; S-AIC model averaging

现实生活中, 数据往往存在不确定性. 有些模型能很好拟合已持有的数据, 但进行预测时却与真实情况相差较大, 误导决策者做出错误的决策^[1]. 模型平均的出现解决了这一问题, 使预测模型变得更加稳定.

传统的变量选择方法代表是: 赤池信息量准则 (Akaike's information criterion, AIC) 由文献[2]在研究信息论尤其是在解决时间序列的定阶问题时提出, 现常被用于处理模型选择, 评估模型的复杂度等属性.

① 收稿日期: 2022-03-17

作者简介: 王梓屹, 硕士研究生, 主要从事纵向数据的研究.

目前经典的模型平均法有两类：频率模型平均法 FMA^[3-4] 和贝叶斯模型平均法 BMA^[5-6]（核心思路是通过后验概率来分配权重）。

文献[7]提出的超高维回归下的模型平均法中，考虑了预测因子数量 p 超过样本容量 n 的高维回归问题，为高维回归问题开发了一个新的模型平均程序，主要研究了该方法下的渐近最优性和模型的权重赋予。文献[8]提出的水手刀模型平均法是一类分位数回归模型平均 (quantile regression model averaging)，为了考虑分位数回归模型中的误差项和解释变量之间的相关性，提出了一种重叠模型平均估计模型，该方法证明了在最小化样本外最终预测误差方面的渐近最优性。

这些研究表明，模型平均预测参数通常会比单一的模型选择方法预测的参数实现更低的估计、预测误差和风险。传统的 SAIC 方法尽管表现较好，但耗时极长。本文提供一种新的模型平均思路用以提升模型的稳定性、准确性和运行速度。

1 广义估计方程中的最小 AIC 模型

假设有 k 个模型，根据文献[2]有其中第 i 个模型中赤池信息量为

$$AIC_i = -2\log(L_i) + 2p_i$$

其中： L_i 为第 i 个模型的似然函数， p_i 为第 i 个模型的参数。

由于本文是纵向数据，并且采用的是广义估计方程用于预测，使用拟似然函数计算赤池信息量过于冗杂，本文通过残差平方项代替似然函数项。将第 i 个模型的赤池信息量变为：

$$AIC_i = 2p_i + N \cdot \log\left(\frac{RSS_i}{N}\right) \quad (1)$$

其中：残差平方和 $RSS_i = \|Y - \hat{Y}_i\|^2$ ； $N = \sum_{i=1}^m n_i$ 为纵向数据中的样本数。在样本数相同的不同广义估计方程中， N 的值都是相等的。

在所有的 k 个模型中，每个模型对应一个赤池信息量。从 k 个赤池信息量中选出赤池信息量最小的模型，将该模型作为预测模型，记为 mAIC (minimum AIC model)。

2 SAIC 模型平均法

文献[3]较为详细提到了 SAIC 模型平均法。其中详细说明了模型平均法的思路和 SAIC 方法的推导，并主要研究了模型的权重赋值。

由赤池信息量的定义

$$AIC_i = -2\log(L_i) + 2p_i$$

我们不难通过变形得到两个拥有不同参数量和似然函数模型的赤池信息量函数比值

$$\frac{L_i \exp(-p_i)}{L_j \exp(-p_j)} = \frac{\exp\left(-\frac{1}{2}AIC_i\right)}{\exp\left(-\frac{1}{2}AIC_j\right)} \quad (2)$$

值得一提的是，在惩罚项相等的时候，该函数会变为似然函数的优势比。

由(2)式可以得到一个关于模型权重 ω_k 的计算公式。假设有 K 个子模型做模型平均，则第 k 个子模型在该模型平均中分配到的权重为

$$\omega_k = \frac{\exp\left(-\frac{1}{2}AIC_k\right)}{\sum_{i=1}^K \exp\left(-\frac{1}{2}AIC_i\right)}, \quad k=1, 2, \dots, K \quad (3)$$

由该方法得出的模型权重有以下特点:

1) 拥有相同赤池信息量的模型具有相同的权重.

2) 假设利用 K 个子模型做模型平均, 这些权重将仅与赤池信息量有关, 和惩罚项无关(他们拥有相同的变量数 p).

3) 拥有更小赤池信息量的模型具有更大的权重.

3 基于 SAIC 的 4 类模型平均法

本文研究的纵向数据结构的参数和符号说明如下:

设 $\{(y_{ij}, x_{ij}); j = 1, 2, \dots, n_i\}$ 是第 i 个个体的观测, 其中 $(i = 1, 2, \dots, m)$, 设置 $\mathbf{y}_i (n_i \times 1) = (y_{i1}, y_{i2}, \dots, y_{in_i})^T$, 且 $\mathbf{X}_i (n_i \times p) = (x_{i1}, x_{i2}, \dots, x_{in_i})$, 则纵向数据为如下线性回归模型

$$\mathbf{y}_i = \mathbf{X}_i \boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (4)$$

其中: \mathbf{y} 为 $(N \times 1)$ 的向量, $\mathbf{y} = (y_1^T, y_2^T, \dots, y_m^T)^T$, \mathbf{X} 为 $(N \times p)$ 维的矩阵, $\mathbf{X} = (\mathbf{X}_1^T, \mathbf{X}_2^T, \dots, \mathbf{X}_m^T)^T$, 残差项 $\boldsymbol{\varepsilon}$ 的维度同 \mathbf{y} 一致, 且服从多元正态分布 $\boldsymbol{\varepsilon} \sim N(0, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}$ 为 $N \times N$ 阶的分块对角矩阵, 即

$$\boldsymbol{\Sigma} = \text{diag}(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \dots, \boldsymbol{\Sigma}_m)$$

通常情况下, 自变量 p 越多, 样本量 N 需要得就越多, 因此我们要求 $N > p$.

3.1 MSAIC

MSAIC 的核心思路是: 在拥有相同参数数量的所有子集中选取 p 个 (p 为数据的协变量数) 赤池信息量最小的模型, 并对这 p 个模型进行加权. 具体步骤如下:

参数数目为 1 的子集为 $(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p)$, 从中选取赤池信息量最小的模型 (\mathbf{X}_{k1}) ; 参数数目为 2 的子集有 $(\mathbf{X}_1, \mathbf{X}_2), (\mathbf{X}_1, \mathbf{X}_3), \dots, (\mathbf{X}_{p-1}, \mathbf{X}_p)$, 从中选取赤池信息量最小的模型 $(\mathbf{X}_{l1}, \mathbf{X}_{l2})$. 重复此动作共 p 次, 挑选出 p 个子集

$$((\mathbf{X}_{k1}), (\mathbf{X}_{l1}, \mathbf{X}_{l2}), \dots, (\mathbf{X}_{r1}, \mathbf{X}_{r2}, \dots, \mathbf{X}_{rp}))$$

基于广义估计方程, p 个子模型共能算出 p 个估计方程: $(f_1(\cdot), f_2(\cdot), \dots, f_p(\cdot))$.

代入数据后可以得出 p 个关于 Y 的拟合值组成的向量:

$$\hat{\mathbf{Y}}_1 = f_1(\mathbf{X}_{k1})$$

$$\hat{\mathbf{Y}}_2 = f_2(\mathbf{X}_{l1}, \mathbf{X}_{l2})$$

...

$$\hat{\mathbf{Y}}_p = f_p(\mathbf{X}_{r1}, \mathbf{X}_{r2}, \dots, \mathbf{X}_{rp})$$

利用 SAIC 加权法, 对这 p 个子集基于 SAIC 方法进行加权. 这 p 个子集可以得到 p 个赤池信息量. 第 i 个模型的赤池信息量变为

$$AIC_i = 2p_i + N \cdot \log\left(\frac{\|\mathbf{Y} - \hat{\mathbf{Y}}_i\|^2}{N}\right) \quad (5)$$

该模型的权重 ω_i 的计算方式

$$\omega_i = \frac{\exp\left(-\frac{AIC_i}{2}\right)}{\sum_{j=1}^p \exp\left(-\frac{AIC_j}{2}\right)} \quad (6)$$

由此可获得 p 个模型的权重 $(\omega_1, \omega_2, \dots, \omega_p)$.

相较传统的 SAIC 方法, 该计算量从约 $p!$ 降低为 $(2^p - 1)$.

在模型(7)中代入数据后能得出 \mathbf{Y} 的估计值 $\hat{\mathbf{Y}}$:

$$\hat{\mathbf{Y}} = \omega_1 f_1(\mathbf{X}_{k_1}) + \omega_2 f_2(\mathbf{X}_{l_1}, \mathbf{X}_{l_2}) + \cdots + \omega_p f_p(\mathbf{X}_{r_1}, \mathbf{X}_{r_2}, \cdots, \mathbf{X}_{r_p}) \quad (7)$$

该模型平均方法仅对传统的 SAIC 模型平均法的最后的加权步骤进行了改进, 计算量仍然巨大, 为解决这一问题, 本文进一步提出了 MOSAIC 以降低模型平均法的计算量.

3.2 MOSAIC

MOSAIC (Minimum Onward Smooth AIC), 是 MSAIC 的改进方法, 借鉴了逐步回归向前法 (Forward Stepwise). MOSAIC 的具体步骤如下:

1) 从参数数量为 1 的子模型中选择拥有最小赤池信息量(共需计算 p 个模型)的子模型 (\mathbf{X}_{k_1}) .

2) 在参数数量为 2 的所有子模型中(共需计算 $(p-1)$ 个模型), 从拥有 (\mathbf{X}_{k_1}) 的子模型中, 选择拥有最小 AIC 的子模型 $(\mathbf{X}_{k_1}, \mathbf{X}_{k_2})$. 重复此行为, 直至将所有参数加入到子模型 $(\mathbf{X}_{k_1}, \mathbf{X}_{k_2}, \cdots, \mathbf{X}_{k_p})$ 中, 挑选出 p 个子集, 基于广义估计方程算出 p 个估计方程:

$$(f_1(\cdot), f_2(\cdot), \cdots, f_p(\cdot))$$

3) 利用 SAIC 加权法, 仅对这 p 个子集基于 SAIC 方法进行加权, 得到 p 个权重 $(\omega_1, \omega_2, \cdots, \omega_p)$. 代入数据后能得出 \mathbf{Y} 的估计值

$$\hat{\mathbf{Y}} = \omega_1 f_1(\mathbf{X}_{k_1}) + \omega_2 f_2(\mathbf{X}_{k_1}, \mathbf{X}_{k_2}) + \cdots + \omega_p f_p(\mathbf{X}_{k_1}, \mathbf{X}_{k_2}, \cdots, \mathbf{X}_{k_p}) \quad (8)$$

该计算量相较本章第一小节提出的 MSAIC 方法的计算量, 从 $(2^p - 1)$ 降低为 $\frac{p(p+1)}{2}$.

该模型平均方法虽然对 MSAIC 进行了改进, 但其计算量仍然较大, 在维度较高的数据中仍难以快速输出结果. 为解决这一问题, 本文进一步提出了 MOOSAIC.

3.3 MOOSAIC

MOOSAIC (Margin Overweighted of Smooth AIC), 是 MOSAIC 的改进方法.

1) 计算参数数量为 1 的所有子模型的赤池信息量(共需计算 p 个模型), 并依据该赤池信息量按从小到大的顺序, 对子模型进行排序

$$(\mathbf{X}_{k_{1(1)}}, \mathbf{X}_{k_{2(2)}}, \cdots, \mathbf{X}_{k_{p(p)}})$$

p 个子模型所示

$$(\mathbf{X}_{k_{1(1)}}, (\mathbf{X}_{k_{1(1)}}, \mathbf{X}_{k_{2(2)}}), \cdots, (\mathbf{X}_{k_{1(1)}}, \mathbf{X}_{k_{2(2)}}, \cdots, \mathbf{X}_{k_{p(p)}}))$$

其中第 i 个模型为

$$(\mathbf{X}_{k_{1(1)}}, \mathbf{X}_{k_{2(2)}}, \cdots, \mathbf{X}_{k_{i(i)}}) \quad (9)$$

基于广义估计方程, p 个子模型能算出 p 个估计方程

$$(f_1(\cdot), f_2(\cdot), \cdots, f_p(\cdot)) \quad (10)$$

基于 SAIC 方法, 对 p 个模型的赤池信息量进行加权, 得到 p 个权重 $(\omega_1, \omega_2, \cdots, \omega_p)$, 代入数据后能得出 \mathbf{Y} 的估计值:

$$\hat{\mathbf{Y}} = \omega_1 f_1(\mathbf{X}_{k_{1(1)}}) + \omega_2 f_2(\mathbf{X}_{k_{1(1)}}, \mathbf{X}_{k_{2(2)}}) + \cdots + \omega_p f_p(\mathbf{X}_{k_{1(1)}}, \mathbf{X}_{k_{2(2)}}, \cdots, \mathbf{X}_{k_{p(p)}}) \quad (11)$$

与 MOSAIC 相比, $\hat{\mathbf{Y}}$ 的计算量从 $\frac{p(p+1)}{2}$ 降低为 $(2p-1)$.

3.4 MO&MOO 加权模型

将 MOSAIC 模型输出的残差平方和与 MOOSAIC 模型输出的残差平方和进行加权, 且权重变化非线性, 可以给出其权重公式:

$$\omega_{mo} = \frac{\exp(\|\mathbf{Y} - \hat{\mathbf{Y}}_{mo}\|^2)}{\exp(\|\mathbf{Y} - \hat{\mathbf{Y}}_v\|^2) + \exp(\|\mathbf{Y} - \hat{\mathbf{Y}}_{mo}\|^2)} \quad (12)$$

$$\omega_{mo} = \frac{\exp(\|\mathbf{Y} - \hat{\mathbf{Y}}_{mo}\|^2)}{\exp(\|\mathbf{Y} - \hat{\mathbf{Y}}_{mo}\|^2) + \exp(\|\mathbf{Y} - \hat{\mathbf{Y}}_{moo}\|^2)} \quad (13)$$

MOSAIC 和 MOOSAIC 的加权模型为

$$\hat{\mathbf{Y}}_{mo\&moo} = \omega_{mo} f_{mo}(\cdot) + \omega_{moo} f_{moo}(\cdot) \quad (14)$$

记为 MO&MOO 模型.

该模型为 MOSAIC 和 MOOSAIC 利用指数进行的简单加权. 引入该模型是为说明: 并非加权模型的部头越大, 模型拟合效果就越好越稳定. 之后的结果, 无论是预测精度还是稳定效果, MO&MOO 模型都远不如它的两个组成部分: MOOSAIC 模型和 MOSAIC 模型.

4 模拟研究

本章借助高维纵向数据的模型平均估计^[9]生成模拟数据, 并在此基础上对模拟数据的各项参数进行调试, 以确保模型预测的稳定性和精准性. 本章模拟研究的抽样方法采用的是 Bootstrap 自助法, 随机挑选训练集和测试集, 并用不同方法进行拟合, 重复该过程 100 次, 之后将对这 100 次拟合得到的预测残差平方和进行研究. 评价模型预测效果的标准是: 预测残差平方和的中位数和四分位差.

4.1 模拟生成数据

生成模拟数据用于实验. 过程如下:

$$y_{ij} = \mu_{ij} + \varepsilon_{ij}, \quad i = 1, 2, \dots, m, \quad j = 1, 2, \dots, n_i \quad (15)$$

其中: $\mu_{ij} = \sum_{k=1}^p x_{ij,k} \beta_k + \varepsilon_{ij}$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n_i$, m 为总个体数, n_i 为每个个体内的观测数, 总样本数为 $\sum_{i=1}^m n_i$; β_k , $k = 1, 2, \dots, p$, 共生成 p 个模拟变量的系数, 其中有 s 个真实变量由均值为 0, 标准差为 0.5 的正态分布生成; 其余虚假变量值设置为 0, 生成的数据是 p 维数据. 解释变量的均值为 0_p . 其中, 多元数据的协方差矩阵为

$$\begin{pmatrix} 1 & d & d^2 & \cdots & d^{p-1} \\ d & 1 & d & \cdots & d^{p-2} \\ \vdots & \vdots & \vdots & & \vdots \\ d^{p-1} & d^{p-2} & d^{p-3} & \cdots & 1 \end{pmatrix}$$

由于本文计算赤池信息量时使用的方法是利用残差平方和代替似然函数, 因此默认残差服从正态分布. 为了研究残差服从正态分布是否会对模型预测产生较大影响, 本文在之后的实验中对残差项不断进行重复实验, 以确保在极端残差情况下仍能保持预测的稳定性.

将随机误差项设为 $c \times AR(1)$, 其中: $AR(1)$ 过程为

$$\varepsilon_{i,j} = \rho \varepsilon_{i,j-1} + v_{i,j} \quad (16)$$

c 为调整随机误差项的一个压缩参数, $v_{i,j}$ 服从正态分布 $N(0, 1)$. 初始 c 值为 1, 初始 ρ 值为 0.5.

考虑如下的具体设计: 令

$$p = 30, \quad m = 500, \quad n_i = 8,$$

$$c \in \{0.2, 1, 2\}, \quad \rho \in \{0.2, 0.4, 0.6\}, \quad d \in \{0.1, 0.5, 0.9\}$$

从总体 500 个个体中挑选 350 个个体作为训练集, 150 个个体作为测试集, 即训练集样本量为 2800, 测试集为 1200.

在实验中，我们将提出的模型平均估计量 MOOSAIC 记为 moo，MOSAIC 与 MOOSAIC 加权拟合模型平均记为 mo&moo，逐步回归法最小 AIC 模型选择记为 mAIC，Scad 模型选择记为 ex_scad，Lasso 模型选择记为 ex_lasso 和全模型，并进行横向比较。通过计算不同模型下的残差平方和来比较各种方法的优劣，重复 100 次实验，每次实验输出 7 个模型的残差平方和，利用残差平方和绘制箱型图来对比各模型的优劣，同时给出了不同情况下预测残差平方和的上四分位数、中位数和下四分位数与四分位差的统计表（表 1—5），其中：1st Qu. 表示下四分位数；Median 表示中位数；3rd Qu. 表示上四分位数；Qu. Deviation 表示四分位差；箱型图中纵坐标为残差平方和。

4.2 模拟实验

实验 1 变动协方差矩阵

保持自回归模型里的常系数 0.2 和压缩参数 1 不变，通过控制自回归模型里的常系数和压缩参数，使得协方差矩阵中的 d 值分别取 0.1, 0.5, 0.9，各方法预测残差平方和箱型图见图 1，预测残差平方和的上四分位数、中位数和下四分位数与四分位差见表 1—3。

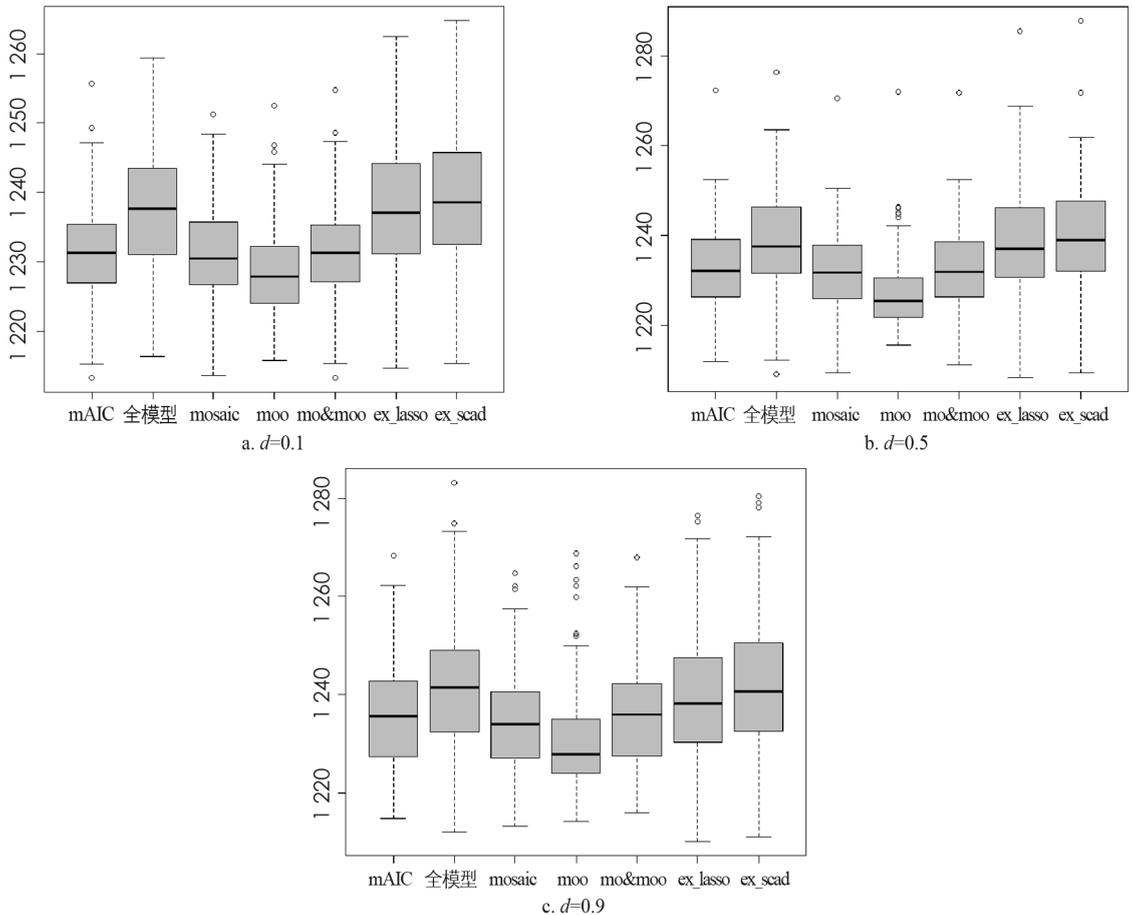


图 1 在 d 取不同值时各方法预测残差平方和箱型图的比较

表 1 $d=0.1$ 时预测残差平方和的上四分位数、中位数和下四分位数与四分位差

	mAIC	全模型	MOSAIC	moo	mo&moo	lasso	scad
1st Qu.	1 226.904	1 231.066	1 226.762	1 224.210	1 227.039	1 231.240	1 232.610
Median	1 231.304	1 237.673	1 230.425	1 227.836	1 231.239	1 237.127	1 238.527
3rd Qu.	1 235.197	1 243.363	1 235.582	1 232.163	1 235.055	1 244.076	1 245.666
Qu. Deviation	8.293	12.297	8.820	7.952	8.016	12.835	13.056

表 2 $d=0.5$ 时预测残差平方和的上四分位数、中位数、下四分位数与四分位差

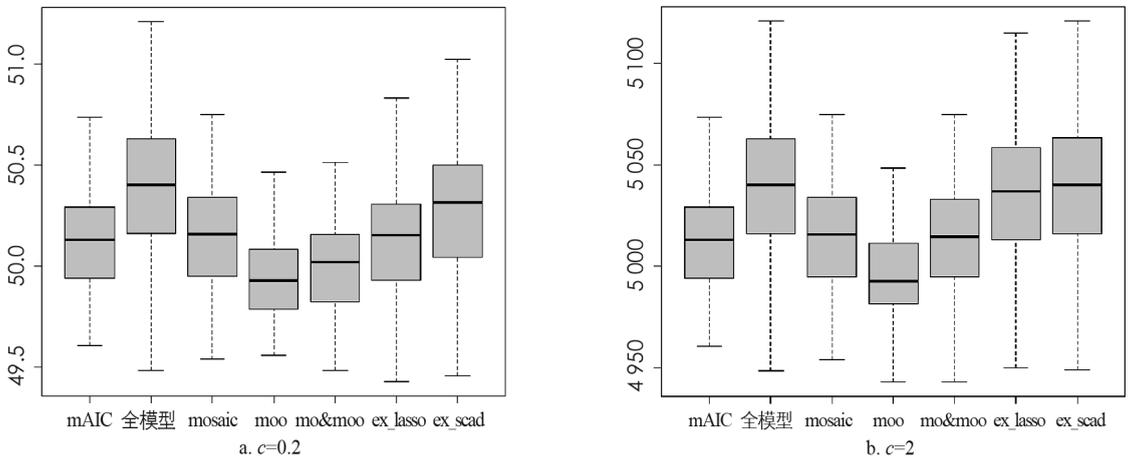
	mAIC	全模型	MOSAIC	moo	mo&moo	lasso	scad
1st Qu.	1 226.491	1 231.675	1 226.097	1 221.828	1 226.481	1 230.988	1 232.376
Median	1 232.121	1 237.679	1 231.799	1 225.447	1 232.067	1 237.101	1 239.064
3rd Qu.	1 239.198	1 246.314	1 238.029	1 230.469	1 238.628	1 246.077	1 247.640
Qu. Deviation	12.708	14.640	11.932	8.641	12.148	15.089	15.264

表 3 $d=0.9$ 时预测残差平方和的上四分位数、中位数、下四分位数与四分位差

	mAIC	全模型	MOSAIC	moo	mo&moo	lasso	scad
1st Qu.	1 227.556	1 232.514	1 227.166	1 223.992	1 227.683	1 230.377	1 232.661
Median	1 235.737	1 241.549	1 234.026	1 227.934	1 236.064	1 238.348	1 240.726
3rd Qu.	1 242.752	1 248.828	1 240.521	1 234.779	1 242.293	1 247.527	1 250.311
Qu. Deviation	15.196	16.314	13.355	10.787	14.609	17.150	17.650

实验 2 变动压缩参数

通过控制协方差矩阵中的 d 值为 0.5 且保持自回归模型里的常系数 0.2 不变,使得压缩参数 c 依次取 0.2 和 2,各方法预测残差平方和箱型图见图 2,预测残差平方和的上四分位数、中位数和下四分位数与四分位差见表 4—5.

图 2 在 c 取不同值时各方法预测残差平方和箱型图的比较表 4 $c=0.2$ 时预测残差平方和的上四分位数、中位数和下四分位数与四分位差

	mAIC	全模型	MOSAIC	moo	mo&moo	lasso	scad
1st Qu.	49.944	50.164	49.951	49.788	49.822	49.931	50.046
Median	50.129	50.401	50.157	49.928	50.018	50.152	50.316
3rd Qu.	50.291	50.631	50.335	50.080	50.152	50.308	50.498
Qu. Deviation	0.346	0.466	0.385	0.292	0.330	0.377	0.452

表 5 $c=2$ 时预测残差平方和的上四分位数、中位数和下四分位数与四分位差

	mAIC	全模型	MOSAIC	moo	mo&moo	lasso	scad
1st Qu.	4 994.437	5 016.417	4 995.089	4 981.373	4 995.089	5 013.150	5 016.412
Median	5 012.944	5 040.125	5 015.702	4 992.345	5 014.385	5 036.924	5 040.138
3rd Qu.	5 029.078	5 063.060	5 033.547	5 010.599	5 033.223	5 058.595	5 063.177
Qu. Deviation	34.641	46.643	38.458	29.226	38.134	45.445	46.765

5 结论

为解决 SAIC 模型在大数据下面临运算时间长的困境, 本文提出了 MSAIC, MOSAIC, MOOSAIC 3 种基于 SAIC 方法的模型平均法. 提升 SAIC 方法下的运行效率、预测精度和预测稳定性所需运行时长从大到小排序依次为 MSAIC, MOSAIC 和 MOOSAIC. 通过数值实验将 3 种方法与多种经典预测方法, 如 Lasso, Scad 等方法进行比较最后结果验证了 MOOSAIC 大幅度提升了运行效率、运行精度和预测稳定性, 且运行时长从数量级 $p!$ 降低为数量级 $(2p-1)$.

参考文献:

- [1] 田文玲. 线性模型平均的一些理论性质研究 [D]. 上海: 华东师范大学, 2019.
- [2] AKAIKE H. A New Look at the Statistical Model Identification [J]. IEEE Transactions on Automatic Control, 1974, 19(6): 716-723.
- [3] HJORT N L, CLAESKENS G. Frequentist Model Average Estimators [J]. Journal of the American Statistical Association, 2003, 98(464): 879-899.
- [4] BUCKLAND S T, BURNHAM K P, AUGUSTIN N H. Model Selection: an Integral Part of Inference [J]. Biometrics, 1997, 53(2): 603.
- [5] DRAPER D. Assessment and Propagation of Model Uncertainty [J]. Journal of the Royal Statistical Society: Series B (Methodological), 1995, 57(1): 45-70.
- [6] RAFTERY A E, MADIGAN D, HOETING J A. Bayesian Model Averaging for Linear Regression Models [J]. Journal of the American Statistical Association, 1997, 92(437): 179-191.
- [7] ANDO T, LI K C. A Model-Averaging Approach for High-Dimensional Regression [J]. Journal of the American Statistical Association, 2014, 109(505): 254-265.
- [8] LU X, SU L J. Jackknife Model Averaging for Quantile Regressions [J]. Journal of Econometrics, 2015, 188(1): 40-58.
- [9] 陈心洁, 赵志豪. 高维纵向数据的模型平均估计 [J]. 系统科学与数学, 2020, 40(7): 1297-1324.

责任编辑 张 枸