重庆理工大学学报(自然科学)
Journal of Chongqing University of Technology(Natural Science)
2021 年第 35 卷第 4 期
Vol. 35 No. 4 2021

doi: 10.3969/j.issn.1674-8425(z).2021.04.029

高维数据下线性模型的序列相关检验

刘 锋,付馨苇,胡天英,陈俊霖

(重庆理工大学 理学院,重庆 404100)

摘　　要: 主要考虑了在高维数据下线性模型的序列相关检验。估计了线性模型的参数,建立了对线性模型随机误差进行序列相关检验的 $V_{T,P}$ 检验统计量,并得到了原假设下检验统计量的渐近分布。通过数值模拟研究了统计量在有限样本下的表现。

关 键 词: 序列相关检验; 高维数据; 线性模型; $V_{T,P}$ 检验

中图分类号: O175.7　　**文献标识码:** A　　**文章编号:** 1674-8425(2021)04-0219-05

Testing Serial Correlation in Linear Model with High Dimensional Data

LIU Feng, FU Xinwei, HU Tianying, CHEN Junlin

(School of Science, Chongqing University of Technology, Chongqing 404100, China)

Abstract: In this paper, we investigate the test for serial correlation in linear model with high dimensional data. First, the parameters of the linear models are estimated. Second, we establish the $V_{T,P}$ statistics to test the serial correlation of random errors in linear models, and further derive the asymptoticdistribution under null hypothesis. Finally, the properties of statistics are studied by the simulation.

Key words: serial correlation test; high-dimensional; linear models; $V_{T,P}$-test

线性模型作为统计学中最基础的模型之一,在过去的几十年里得到了大量运用。由于其结构简明、理论健全,因而在金融、自动化技术、电力工业、医药卫生和航空航天等领域均有广泛涉及。在线性模型大量运用的同时,对于模型中残差是否存在序列相关的检验也成为数据分析中的重要工作。如果残差是序列相关的,则得到的最小二乘估计不是有效的,更糟糕的是,若存在很强的序列相关性则意味着忽略了一些重要的解释变量,导致模型被误用的可能。对于本文所需进行的序列相关检验,参考 HU 等[1]所提出的 $V_{T,P}$ 检验统计量,在原假设下得到 $V_{T,P}$ 检验统计量的渐近分布。本文研究在高维数据下的线性模型序列相关问题。在计算机和信息技术高速发展的今天,随着积累的数据越来越多,高维数据涉及范围越来越广,加之许多低维数据的经典处理方法,如主成分

收稿日期: 2020-09-04
基金项目: 国家自然科学基金项目(11871124)
作者简介: 刘锋,男,博士,副教授,主要从事非参数统计研究,E-mail: fliu@cqut.edu.cn。

本文引用格式: 刘锋,付馨苇,胡天英,等. 高维数据下线性模型的序列相关检验[J]. 重庆理工大学学报(自然科学),2021,35(4):219-223.

Citation format: LIU Feng, FU Xinwei, HU Tianying, et al. Testing Serial Correlation in Linear Model with High Dimensional Data[J]. Journal of Chongqing University of Technology(Natural Science),2021,35(4):219-223.

分析、回归分析、聚类分析和 AHP 等在处理高维数据时有着相当大的困难,进而高维统计方法的研究在各个领域变得十分重要。因此,相较于以往序列相关检验针对的低维数据,针对高维情形进行的检验将能更好地解决新型实际问题。所谓高维数据,是指协变量维数 p 大于样本量 n 的数据,也正因为数据的高维从而不可避免地产生“维数灾祸”问题。为解决维数带来的“灾难”,许多学者提出了各种不同的方法和算法,包括但不限于:① 数据的聚类[2],包含 Kohonen 自组织特征映射[3-4]、多维缩放[5]、基于分形的降维[6-7];② 变量的筛选,包括 Tibshirani(1996) 提出的 LASSO (least absolute shrinkage and selection operator) 方法[8],Yang 等[9]将 Dantzig 变量选择应用到高维部分线性模型,李冰月[10]利用 profile 最小二乘方法结合 RAR 部分惩罚方法对超高维部分线性模型进行变量估计等,进一步地,对于本文针对的一般线性模型,LASSO 类方法还可以推广到自适应 LASSO[11]、松弛 LASSO[12] 以及 Group LASSO[13] 等;③ 利用方差分析和因子分析进行降维[14];④ 对高维数据进行特征降维[15]等。在探寻降维算法的同时,还有一些学者考虑了降维对原始数据的影响[16]。

本文采用的检验方法与传统 D-W 检验不同,其可检验高阶序列相关且无需对误差的分布做出任何假定,同时结合 L1 惩罚函数的变量估计对高维数据下线性模型的序列相关进行检验。

1 方法及主要结果

记 (Y_i, X_i), $i=1,2,\cdots,n$ 为来自模型(1)的 n 个 i. i. d 样本。

$$Y_i = X_i^{\mathrm{T}}\beta + \varepsilon_i \tag{1}$$

式中:Y_i 为响应变量,X_i 是 p 维的随机向量且 $X_i \sim N(0,\Sigma)$, $\varepsilon_i \sim N(0,\sigma^2)$。在本文中,考虑 $p > n \to \infty$。

假设 ε_i 是来自于模型 AR(k) 的随机误差:

$$\varepsilon_i = a_1\varepsilon_{i-1} + a_2\varepsilon_{i-2} + \cdots a_k\varepsilon_{i-k} + e_i$$

或者为模型 MA(k) 的随机误差:

$$\varepsilon_i = a_1 e_{i-1} + a_2 e_{i-2} + \cdots a_k e_{i-k} + e_i$$

式中:e_i 为 i. i. d 的随机变量,同时有 $E(e_i)=0$, $\mathrm{Var}(e_i) = \sigma^2 < \infty$。对于 a_i, $i=1,2,\cdots,k$,其为未知的回归系数,且对于模型 AR(k) 是固定的。

针对序列相关的检验是在模型(1)的基础上对误差序列 ε_i 进行的,即只需要检验 AR(k) 模型亦或是 MA(k) 模型中的系数是否全为零即可。若模型系数全为零,则表明(1)的误差序列不存在序列相关,反之则存在高阶序列相关。

令 $a = (a_1, a_2, \cdots, a_k)^{\mathrm{T}}$,因此检验为

$$\mathrm{H}_0: a = 0 \leftrightarrow \mathrm{H}_1: a \neq 0$$

令 $u_j = E(\varepsilon_i\varepsilon_{i+j})$ 以及 $U = (u_1, u_2, \cdots, u_k)$, $j = 1,2,\cdots,k$。此时关于该检验的原假设和备择假设为:

$$\mathrm{H}_0: U = 0 \leftrightarrow \mathrm{H}_1: U \neq 0$$

令

$$\begin{cases} z_{i1} = \varepsilon_i\varepsilon_{i+1} = (Y_i - X_i^{\mathrm{T}}\beta)(Y_{i+1} - X_{i+1}^{\mathrm{T}}\beta) \\ \vdots \\ z_{ik} = \varepsilon_i\varepsilon_{i+k} = (Y_i - X_i^{\mathrm{T}}\beta)(Y_{i+k} - X_{i+k}^{\mathrm{T}}\beta) \end{cases}$$

并记作

$$Z_i = (z_{i1}, z_{i2}, \cdots, z_{ik})^{\mathrm{T}}, \ i = 1,2,\cdots,N-k$$

在原假设下有 $E(Z_i) = 0$, $i = 1,2,\cdots,N-k$,且 $E(Z_i) = 0$ 意味着 $\{\varepsilon\}_i^N$ 中没有序列相关性。因此,对于序列相关性的检验等同于检验 $E(Z_i) = 0$, i. e.,

$$\mathrm{H}_0: E(Z_i) = 0 \leftrightarrow \mathrm{H}_1: E(Z_i) \neq 0$$

参考 HU[1],可以利用 $V_{T,P}$ 检验统计量进行检验。对任意的 p 维向量 $\boldsymbol{a}$,记 $\|a\|_1 = \sum_{i=1}^{p}|a_i|$。记 $\boldsymbol{Y} = (Y_1, Y_2, \cdots, Y_n)^{\mathrm{T}}$, $\boldsymbol{X}^{\mathrm{T}} = (X_1, X_2, \cdots, X_n)$, $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \cdots \varepsilon_n)^{\mathrm{T}}$, $\hat{\Sigma} = (1/n)\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$,基于 L_1 惩罚函数,定义 β 的估计为

$$\hat{\beta} \equiv \hat{\beta}(\lambda) = \arg\min_{\beta}\left\{\frac{1}{2n}\|\boldsymbol{Y} - \boldsymbol{X}\beta\|_2^2 + \lambda\|\beta\|_1\right\} \tag{2}$$

其中 $\lambda \geqslant 0$ 为 $L1$ 正则化参数。特别地,令 $T = n - k$,于是在原假设的约束下,有 $V_{T,P}$ 检验统计量

$$V_{T,P} = \frac{1}{\sqrt{T}}\sum_{i=1}^{T} Z_i \tag{3}$$

将 β 替换为 $\hat{\beta}$,相应地 Z_i 替换为 $\hat{Z}_i$,则有估计的 $V_{T,P}$ 检验统计量

$$\hat{V}_{T,P} = \frac{1}{\sqrt{T}}\sum_{i=1}^{T} \hat{Z}_i \tag{4}$$

为得到主要结论，需要如下假设条件：

条件1　$E(\varepsilon_i^4) < \infty$

条件2　记 $\lambda \cong \sqrt{\log p/n} \to 0$，则有 $\|\hat{\beta}-\beta\|_1 = O_p(s_0\lambda)$（见 peter J. Bickel(2009)定理7.2）。本文中 $s_0 \cong (\sqrt{n}/\log p)\,o(1)$

定理1　在条件1和条件2以及零假设下，当 $T\to\infty$ 时，有

$$\hat{V}_{T,P} \xrightarrow{L} N(0, \sigma^4 \boldsymbol{I}_k)$$

其中 $\boldsymbol{I}_k$ 为 $k\times k$ 的单位阵。

在定理1下便得到了原假设中检验统计量的渐近分布，因此当利用统计推断时，在已知的显著性水平 α 下，可以在检验统计量的值大于渐近分布的 $1-\alpha$ 分位点时，拒绝原假设，即认为模型(1)的误差序列存在序列相关性，反之同理。

2　数值模拟

通过相当部分的数值模拟研究高维数据下线性模型的序列相关性检验。考虑模型 $Y_i = X_i^{\mathrm{T}}\beta + \varepsilon_i$，设 X_i 产生于正态分布 $N(0,\Sigma)$，利用10折交叉验证对 β 的相关参数进行选择，选出最优参数后再利用 $L1$ 惩罚函数对 β 进行估计。对于 ε_i，则分别假定其服从以下模型：

1) AR(1)模型：$\varepsilon_i = a_1\varepsilon_{i-1} + e_i$

2) AR(2)模型：$\varepsilon_i = a_1\varepsilon_{i-1} + a_2\varepsilon_{i-2} + e_i$

3) MA(1)模型：$\varepsilon_i = a_1 e_{i-1} + e_i$

4) MA(2)模型：$\varepsilon_i = a_1 e_{i-1} + a_2 e_{i-2} + e_i$

同时假定误差 e_i 服从正态分布。取显著性水平 $\alpha = 0.05$，样本量 n 取200、400和600，维数 p 分别取400、600和800时，各做1 000次模拟，有表1～12。

表1　AR(1)模型　$n=200$

a_1	0	0.2	0.4	0.6	0.8
$p=400$	0.039	0.588	0.989	0.998	0.996
$p=600$	0.033	0.585	0.967	0.991	0.999
$p=800$	0.031	0.532	0.978	0.988	0.991

表2　AR(1)模型　$n=400$

a_1	0	0.2	0.4	0.6	0.8
$p=400$	0.049	0.952	1.000	1.000	1.000
$p=600$	0.044	0.941	1.000	1.000	1.000
$p=800$	0.045	0.946	1.000	1.000	1.000

表3　AR(1)模型　$n=600$

a_1	0	0.2	0.4	0.6	0.8
$p=400$	0.039	0.994	1.000	1.000	1.000
$p=600$	0.038	0.991	1.000	1.000	1.000
$p=800$	0.036	0.997	1.000	1.000	1.000

表4　MA(1)模型　$n=200$

a_1	0	0.2	0.4	0.6	0.8
$p=400$	0.040	0.540	0.969	0.996	0.996
$p=600$	0.026	0.514	0.947	0.99	0.991
$p=800$	0.022	0.509	0.940	0.984	0.979

表5　MA(1)模型　$n=400$

a_1	0	0.2	0.4	0.6	0.8
$p=400$	0.047	0.932	1.000	1.000	1.000
$p=600$	0.040	0.928	1.000	1.000	1.000
$p=800$	0.034	0.922	1.000	1.000	1.000

表6　MA(1)模型　$n=600$

a_1	0	0.2	0.4	0.6	0.8
$p=400$	0.043	0.995	1.000	1.000	1.000
$p=600$	0.043	0.996	1.000	1.000	1.000
$p=800$	0.026	0.994	1.000	1.000	1.000

表7　AR(2)模型　$n=200$

$a=(a_1,a_2)$	(0，0)	(0，0.5)	(0.3，0.4)	(0.4，−0.6)	(−0.2，0.6)
$p=400$	0.040	1.000	1.000	0.951	0.994
$p=600$	0.034	1.000	1.000	0.957	0.992
$p=800$	0.036	1.000	1.000	0.939	0.995

表8　AR(2)模型　$n=400$

$a=(a_1,a_2)$	(0，0)	(0，0.5)	(0.3，0.4)	(0.4，−0.6)	(−0.2，0.6)
$p=400$	0.049	1.000	1.000	0.935	0.992
$p=600$	0.043	1.000	1.000	0.942	0.993
$p=800$	0.044	1.000	1.000	0.951	0.995

表 9 AR(2) 模型 $n=600$

$a=(a_1,a_2)$	(0, 0)	(0, 0.5)	(0.3, 0.4)	(0.4, −0.6)	(−0.2, 0.6)
$p=400$	0.036	1.000	1.000	0.944	0.993
$p=600$	0.053	1.000	1.000	0.937	0.992
$p=800$	0.047	1.000	1.000	0.942	0.995

表 10 MA(2) 模型 $n=200$

$a=(a_1,a_2)$	(0, 0)	(0, 0.5)	(0.3, 0.4)	(0.4, −0.6)	(−0.2, 0.6)
$p=400$	0.045	1.000	1.000	0.889	0.999
$p=600$	0.041	1.000	1.000	0.890	0.998
$p=800$	0.040	1.000	1.000	0.867	0.997

表 11 MA(2) 模型 $n=400$

$a=(a_1,a_2)$	(0, 0)	(0, 0.5)	(0.3, 0.4)	(0.4, −0.6)	(−0.2, 0.6)
$p=400$	0.047	1.000	1.000	0.893	0.998
$p=600$	0.043	1.000	1.000	0.895	1.000
$p=800$	0.040	1.000	1.000	0.869	0.999

表 12 MA(2) 模型 $n=600$

$a=(a_1,a_2)$	(0, 0)	(0, 0.5)	(0.3, 0.4)	(0.4, −0.6)	(−0.2, 0.6)
$p=400$	0.048	1.000	1.000	0.885	0.997
$p=600$	0.052	1.000	1.000	0.859	0.999
$p=800$	0.033	1.000	1.000	0.877	0.999

从上述结果中可以看到,当误差服从正态分布时,若模型不存在序列相关,则检验结果在显著性水平 0.05 附近波动,且当固定样本量 n,检验统计量的 size 随着维数 p 的增大而大致呈减小趋势,说明当样本量固定不变时,维数的增大将导致检验统计量的 size 效果变差。而在备择假设中,无论是 AR 模型还是 MA 模型,当相关系数变大时,检验统计量的 power 均趋向于 1。

3 定理的证明

对于任意的 $m\times n$ 矩阵 $\boldsymbol{A}$,定义 $\|A\|_\infty = \max\limits_{1\leqslant i\leqslant m,1\leqslant j\leqslant n}|a_{i,j}|$。

引理 1 在概率至少为 $1-Cp^{-1}$ 时

$$\left\|\frac{1}{n}\sum_{i=1}^{T}\boldsymbol{X}_{i+k}\boldsymbol{X}_i^{\mathrm{T}}\right\|_\infty \leqslant C\sqrt{\frac{T\log p}{n}} \tag{5}$$

证明: 证明见 Chang et al. (2010)[17] 的引理 2 和引理 3。

定理 2 在条件 1、条件 2 及原假设下有

$$\frac{1}{\sqrt{T}}\sum_{i=1}^{T}\hat{Z}_i \xrightarrow{L} N(0,\sigma^4 I_k)\ \text{as}\ n\to\infty \tag{6}$$

式中 I_k 为 $k\times k$ 的单位阵。

证明: 首先考虑第 j 项下的 $\hat{Z}_i$, i.e.,

$$\frac{1}{\sqrt{T}}\sum_{i=1}^{T}\hat{Z}_{ij} = \frac{1}{\sqrt{T}}\sum_{i=1}^{T}\hat{\varepsilon}_i\hat{\varepsilon}_{i+j} =$$

$$\frac{1}{\sqrt{T}}\sum_{i=1}^{T}[Y_i - \boldsymbol{X}_i^{\mathrm{T}}\hat{\beta}][Y_{i+j} - \boldsymbol{X}_{i+j}^{\mathrm{T}}\hat{\beta}] =$$

$$\frac{1}{\sqrt{T}}\sum_{i=1}^{T}[\varepsilon_i - \boldsymbol{X}_i^{\mathrm{T}}(\hat{\beta}-\beta)][\varepsilon_{i+j} - \boldsymbol{X}_{i+j}^{\mathrm{T}}(\hat{\beta}-\beta)] =$$

$$\frac{1}{\sqrt{T}}\sum_{i=1}^{T}[\varepsilon_i\varepsilon_{i+j} - \varepsilon_i\boldsymbol{X}_{i+j}^{\mathrm{T}}(\hat{\beta}-\beta) -$$

$$\varepsilon_{i+j}\boldsymbol{X}_i^{\mathrm{T}}(\hat{\beta}-\beta) + (\hat{\beta}-\beta)^{\mathrm{T}}\boldsymbol{X}_i\boldsymbol{X}_{i+j}^{\mathrm{T}}(\hat{\beta}-\beta)] \equiv$$

$$\frac{1}{\sqrt{T}}\sum_{i=1}^{T}\varepsilon_i\varepsilon_{i+j} - \Delta_1 - \Delta_2 + \Delta_3$$

令 $\xi_j = \max\{\|\boldsymbol{X}_{1+j,s}\varepsilon_1\|_{\psi_1}, s\leqslant p\}$ 对于 Δ_1,利用中心下指数随机变量结合 Bernstein-type 不等式(Vershynin (2010)[18]),对于任意的 $t\geqslant 0$,均有

$$P\left(\left\|\frac{1}{T}\sum_{i=1}^{T}\varepsilon_i\boldsymbol{X}_{i+j}\right\|_\infty > t\right) = P\left(\left\|\sum_{i=1}^{T}\varepsilon_i\boldsymbol{X}_{i+j}\right\|_\infty > Tt\right)\leqslant$$

$$P\left(\bigcup_{s=1}^{p}\left\{\left|\sum_{i=1}^{T}\varepsilon_i\boldsymbol{X}_{i+j,s}\right| > Tt\right\}\right)\leqslant$$

$$\sum_{s=1}^{p}P\left(\left\{\left|\sum_{i=1}^{T}\varepsilon_i\boldsymbol{X}_{i+j,s}\right| > Tt\right\}\right)\leqslant 2p\exp\left(-C\frac{Tt^2}{\xi_j^2}\right) =$$

$$2\exp\left(-C\frac{Tt^2}{\xi_j^2} + \log p\right)$$

当 $t^2 = 2\dfrac{\xi_j^2}{C}\dfrac{\log p}{T} = 2\exp(-2\log p + \log p) = 2/p \to 0$

因此,当 $t\lesssim\sqrt{\log p/T}$ 时,有

$$P\left(\left\|\frac{1}{T}\sum_{i=1}^{T}\varepsilon_i\boldsymbol{X}_{i+j}\right\|_\infty > C\sqrt{\log p/n}\right)\to 0$$

即 $\left\|\dfrac{1}{T}\sum\limits_{i=1}^{T}\varepsilon_i\boldsymbol{X}_{i+j}\right\|_\infty = O_p(\sqrt{\log p/T})$,则对于 Δ_1 有

$$\Delta_1 = \left|\frac{1}{\sqrt{T}}\sum_{i=1}^{T}\varepsilon_i\boldsymbol{X}_{i+j}^{\mathrm{T}}(\hat{\beta}-\beta)\right| =$$

$$\sqrt{T}\left|\frac{1}{T}\sum_{i=1}^{T}\varepsilon_i\boldsymbol{X}_{i+j}^{\mathrm{T}}(\hat{\beta}-\beta)\right| \leqslant$$

$$\sqrt{T}\left\|\frac{1}{T}\sum_{i=1}^{T}\varepsilon_i\boldsymbol{X}_{i+j}\right\|_{\infty}\|\hat{\beta}-\beta\|_1 =$$

$$\sqrt{n}O_p(\sqrt{\log p/T})O_p(s_0\lambda) = o_p(1)$$

同时在 Δ_3 中有

$$\Delta_3 = \left|\frac{1}{\sqrt{T}}\sum_{i=1}^{T}(\hat{\beta}-\beta)^{\mathrm{T}}\boldsymbol{X}_i\boldsymbol{X}_{i+j}^{\mathrm{T}}(\hat{\beta}-\beta)\right| \leqslant$$

$$\frac{n}{\sqrt{T}}\left\|\frac{1}{n}\sum_{i=1}^{T}\boldsymbol{X}_{i+j}\boldsymbol{X}_i^{\mathrm{T}}\right\|_{\infty}\|\hat{\beta}-\beta\|_1^2 =$$

$$\frac{n}{\sqrt{T}}O_p(\sqrt{\log p/n})O_p(s_0^2\lambda^2) = o_p(1)$$

至此得到

$$\frac{1}{\sqrt{T}}\sum_{i=1}^{T}\hat{Z}_{ij} = \frac{1}{\sqrt{T}}\sum_{i=1}^{T}\varepsilon_i\varepsilon_{i+j} + o_p(1)$$

从而有

$$\frac{1}{\sqrt{T}}\sum_{i=1}^{T}\hat{Z}_i = \frac{1}{\sqrt{T}}\sum_{i=1}^{T}Z_i + o_p(1)$$

由 Gramer-Wold 方法，根据 m 步相依随机变量中心极限定理可得 $\frac{1}{\sqrt{T}}\sum_{i=1}^{T}\hat{Z}_i \xrightarrow{L} N(0,\sigma^4 I_k)$

定理得证。

参考文献：

[1] HU Xuemei，LIU Feng，WANG Zhizhong. Testing serial correlation in semiparametric varying coefficient partial linear Errors-in-Variables models [J]. Syst Sci & Complexity，2006，22：483 – 494.

[2] LIU M，WANG X L，LIU Y C. A fast clustering algorithm for large-scale and high dimensional data [J]. Acta Automatica Sinica，2009，35(7)：859 – 866.

[3] KOHONEN T. Self-organizing maps [M]. New York：Springer，2001.

[4] KOHONEN T. Self-organization and associative memory [M]. Springer，1984.

[5] THOMAS L GRIFFITHS，MICHAEL L，et al. A multidimensional scaling approach to mental multiplication [J]. Memory and Cognition，2002，30(1)：97 – 106.

[6] AMASTRA F，VINCIARELLI A. Estimating the intrinsic dimension of data with a fractal-based method [J]. IEEE Trans on pattern Ana，2002，24(10)：1404 – 1407.

[7] FRANCESCO C. Data dimensionality estimation methods：A survey [J]. Pattern Recognition，2003，36：2945 – 2954.

[8] JELLE J GOEMAN. L1 penalized estimation in the cox proportional hazards model [J]. Biometrical Journal，2010，52(1)：70 – 84.

[9] YANG Yiping，XUE Liugen. Variable selection and estimation in high-dimensional partially linear models [J]. Chinese Journal of Applied Probability and Statistics，2011，27(2)：172 – 182.

[10] LEE Binyue. Variable selection for ultrahigh dimensional partially linear models [D]. Beijing：Beijing University of Technology，2017.

[11] ZOU Hui. The adaptive lasso and its oracle properties [J]. Journal of the American Statistical Association，2006，101(476)：1418 – 1429.

[12] MEINSHAUSEN N. Relaxed lasso [J]. Computational Statistics and Data Analysis，2006，52(1)：374 – 393.

[13] YUAN Ming，LIN Yi. Model selection and estimation in regression with grouped variables [J]. Journal of the Royal Statistical Society：Series B(Statistical Methodology)，2006，68(1)：49 – 67.

[14] CARSTEN S，GERHARD A，ALEXANDRA S. Analysis of variance and factor analysis for the reduction of high dimensional variation in time series of energy consumption [J]. Allgemeines Statistisches Archiv，2005，89(4)：403 – 418.

[15] HU Jie. Survey on feature dimension reduction for high-dimensional data [J]. Application Research of Computers，2008(9)：2601 – 2606.

[16] AGGARWAL C C. On the effects of dimensionality reduction on high dimensional similarity search [J]. H.2.8 [Database Management]：Database Applications. 2002. DOI：10.1145/375551.383213.

[17] CHANG Jinyuan，YAO Qiwei，ZHOU Wen. Testing for high-dimensional white noise using maximum cross-correlations [J]. Biometrika Trust，2017，104(1)：111 – 127.

[18] ROMAN V，ELDAR Y，KUTYNIOK G. Introduction to the non-asymptotic analysis of random matrices [J]. Theory and Applications，2012(5)：210 – 268.

（责任编辑　林　芳）