附录1

附录1展示定理1的详细证明过程。

该定理的证明过程与文献 Wang 等（2021）中定理2的证明过程类似。具体地，证明定理1中 $P(\hat{r}=r_0)\to 1$ 成立，等价于分别证明 $P(\hat{r}>r_0)\to 0$ 和 $P(\hat{r}<r_0)\to 0$ 成立。

对于真实维数 r_0，易知该维数下模型参数估计值的残差平方和为

$$RSS_i(r_0)=\|\hat{\boldsymbol{z}}_{i,0}-\widehat{\boldsymbol{M}}_{i,0}\hat{\boldsymbol{\theta}}_{i,0}\|_2^2,\qquad i=1,\cdots,p,$$

定义矩阵 $\boldsymbol{H}_{i,0}=\widehat{\boldsymbol{M}}_{i,0}(\widehat{\boldsymbol{M}}_{i,0}\boldsymbol{W}^{-1}\widehat{\boldsymbol{M}}_{i,0})^{-1}\widehat{\boldsymbol{M}}_{i,0}\boldsymbol{W}^{-1}$，根据2范数的相容性和矩阵 $\boldsymbol{H}_{i,0}$ 的幂等性质，可知

$$\begin{aligned}RSS_i(r_0)&=\|(\boldsymbol{I}_{r_0p}-\boldsymbol{H}_{i,0})\frac{1}{n-1}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-1}^{b}\tilde{\varepsilon}_{i,t}\|_2^2\\&\leqslant\|\boldsymbol{I}_{r_0p}-\boldsymbol{H}_{i,0}\|_2^2\times\|\frac{1}{n-1}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-1}^{b}\tilde{\varepsilon}_{i,t}\|_2^2\\&\leqslant\|\frac{1}{n-1}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-1}^{b}\tilde{\varepsilon}_{i,t}\|_2^2\\&=O_p(\frac{1}{n}).\end{aligned}\tag{A1}$$

当 $r>r_0$ 时，设 $\hat{\boldsymbol{z}}_i=(\hat{\boldsymbol{z}}_{i,0},\hat{\boldsymbol{z}}_{i,1})$ 和 $\widehat{\boldsymbol{M}}_i=(\widehat{\boldsymbol{M}}_{i,0},\widehat{\boldsymbol{M}}_{i,1})$ 。另外，矩阵 $\widehat{\boldsymbol{M}}_{i,0}$ 和 $\widehat{\boldsymbol{M}}_{i,1}$ 对应的参数向量分别为 $\hat{\boldsymbol{\theta}}_{i,0}$ 和 $\hat{\boldsymbol{\theta}}_{i,1}$，此时的残差平方和满足

$$\begin{aligned}RSS_i(r)&=\|\hat{\boldsymbol{z}}_i-\widehat{\boldsymbol{M}}_i\hat{\boldsymbol{\theta}}_i\|_2^2\\&=\|\hat{\boldsymbol{z}}_{i,0}-\widehat{\boldsymbol{M}}_{i,0}\hat{\boldsymbol{\theta}}_{i,0}\|_2^2+\|\hat{\boldsymbol{z}}_{i,1}-\widehat{\boldsymbol{M}}_{i,1}\hat{\boldsymbol{\theta}}_{i,1}\|_2^2\\&=O_p(\frac{1}{n}).\end{aligned}\tag{A2}$$

类似地，当 $r<r_0$ 时，设 $\hat{\boldsymbol{z}}_{i,0}=(\hat{\boldsymbol{z}}_i\ ,\hat{\boldsymbol{z}}_{i,2})$ 和 $\widehat{\boldsymbol{M}}_{i,0}=(\widehat{\boldsymbol{M}}_i\ ,\widehat{\boldsymbol{M}}_{i,2})$ ，另外，矩阵 $\widehat{\boldsymbol{M}}_{i,0}$ 和 $\widehat{\boldsymbol{M}}_{i,2}$ 对应的参数向量分别为 $\hat{\boldsymbol{\theta}}_{i,0}$ 和 $\hat{\boldsymbol{\theta}}_{i,2}$，此时的残差平方和有

$$\begin{aligned}RSS_i(r)&=\|\hat{\boldsymbol{z}}_i-\widehat{\boldsymbol{M}}_i\hat{\boldsymbol{\theta}}_i\|_2^2\\&=\|\hat{\boldsymbol{z}}_{i,0}-\widehat{\boldsymbol{M}}_{i,0}\hat{\boldsymbol{\theta}}_{i,0}\|_2^2+\|\hat{\boldsymbol{z}}_{i,1}-\widehat{\boldsymbol{M}}_{i,2}\hat{\boldsymbol{\theta}}_{i,2}\|_2^2\\&=O_p(\frac{1}{n}).\end{aligned}\tag{A3}$$

根据文中采用的 RSS 比率值易知，事件 $r>r_0$ 意味着事件

$$\Psi_1:=\left\{\frac{\max\dfrac{\log RSS(\hat{r}+1)}{\log RSS(\hat{r})+c_n}}{\dfrac{\log RSS(\widehat{r_0}+1)}{\log RSS(\widehat{r_0})+c_n}}>1\right\}$$

成立。而根据式（A1）和式（A2）可知，当 $n\to\infty$ 时，有

$$\frac{\max\limits_{\hat{r}>r_0}\dfrac{\log RSS_i(\hat{r}+1)}{\log RSS_i(\hat{r})+c_n}}{\dfrac{\log RSS_i(\widehat{r_0}+1)}{\log RSS_i(\widehat{r_0})+c_n}}=\frac{\log\dfrac{1}{n}/(c_n+\log\dfrac{1}{n})}{\log\dfrac{1}{n}/(c_n+\log\dfrac{1}{n})}=1,\qquad i=1,\cdots,p,$$

意味着当$n\to\infty$时有$P(\hat{r}>r_0)\to 0$。类似地，根据式（A1）和式（A3）易知当$n\to\infty$时有$P(\hat{r}<r_0)\to 0$。证毕。

附录2

附录2展示定理2的详细证明过程.

根据正文中的模型参数估计方法可知$\widehat{\boldsymbol{\theta}}_i=(\widehat{\boldsymbol{M}}_i\ \widehat{\boldsymbol{W}}^{-1}\widehat{\boldsymbol{M}}_i)^{-1}\widehat{\boldsymbol{M}}_i\ \widehat{W}^{-1}\widehat{\boldsymbol{z}}_i$，故

$$\begin{aligned}
\widehat{\boldsymbol{\theta}}_i-\boldsymbol{\theta}_i&=(\widehat{\boldsymbol{M}}_i\ \widehat{\boldsymbol{W}}^{-1}\widehat{\boldsymbol{M}}_i)^{-1}\widehat{\boldsymbol{M}}_i\ \widehat{\boldsymbol{W}}^{-1}\begin{bmatrix}\dfrac{1}{n}\sum\limits_{t=2}^{n}\widetilde{\boldsymbol{Y}}^a_{t-2}\tilde{Y}^a_{i,t}\\ \vdots\\ \dfrac{1}{n}\sum\limits_{t=2}^{n}\widetilde{\boldsymbol{Y}}^a_{t-r-1}\tilde{Y}^a_{i,t}\end{bmatrix}\\
&=(\widehat{\boldsymbol{M}}_i\ \widehat{\boldsymbol{W}}^{-1}\widehat{\boldsymbol{M}}_i)^{-1}\widehat{\boldsymbol{M}}_i\ \widehat{\boldsymbol{W}}^{-1}\begin{bmatrix}\dfrac{1}{n}\sum\limits_{t=2}^{n}\widetilde{\boldsymbol{Y}}^a_{t-2}(\widetilde{\boldsymbol{Y}}^a_t\boldsymbol{a}_i+\widetilde{\boldsymbol{Y}}^b_{t-1}\boldsymbol{b}_i+\tilde{\varepsilon}_{i,t})\\ \vdots\\ \dfrac{1}{n}\sum\limits_{t=2}^{n}\widetilde{\boldsymbol{Y}}^a_{t-r-1}(\widetilde{\boldsymbol{Y}}^a_t\boldsymbol{a}_i+\widetilde{\boldsymbol{Y}}^b_{t-1}\boldsymbol{b}_i+\tilde{\varepsilon}_{i,t})\end{bmatrix}\\
&=(\widehat{\boldsymbol{M}}_i\ \widehat{\boldsymbol{W}}^{-1}\widehat{\boldsymbol{M}}_i)^{-1}\widehat{\boldsymbol{M}}_i\ \widehat{\boldsymbol{W}}^{-1}\begin{bmatrix}\dfrac{1}{n}\sum\limits_{t=2}^{n}\widetilde{\boldsymbol{Y}}^a_{t-2}\tilde{\varepsilon}_{i,t}\\ \vdots\\ \dfrac{1}{n}\sum\limits_{t=2}^{n}\widetilde{\boldsymbol{Y}}^a_{t-r-1}\tilde{\varepsilon}_{i,t}\end{bmatrix},
\end{aligned}$$

因此定理2的证明可等价于证明以下两个部分：

$$(*)\quad \sqrt{n}\boldsymbol{U}_i^{-\frac{1}{2}}\widehat{\boldsymbol{M}}_i\ \widehat{\boldsymbol{W}}^{-1}\cdot\begin{bmatrix}\dfrac{1}{n}\sum\limits_{t=2}^{n}\widetilde{\boldsymbol{Y}}^a_{t-2}\tilde{\varepsilon}_{i,t}\\ \vdots\\ \dfrac{1}{n}\sum\limits_{t=2}^{n}\widetilde{\boldsymbol{Y}}^a_{t-r-1}\tilde{\varepsilon}_{i,t}\end{bmatrix}\xrightarrow{d}N(\boldsymbol{0},\boldsymbol{I}_{2p-1}),$$

$$(**)\quad \boldsymbol{K}_0(\widehat{\boldsymbol{M}}_i\ \widehat{\boldsymbol{W}}^{-1}\widehat{\boldsymbol{M}}_i)^{-1}\xrightarrow{\mathrm{p}}\boldsymbol{I}_{2p-1}.$$

对于（*）这一部分而言，需证明对于任意非零向量$\boldsymbol{a}=(\boldsymbol{a}_0,\boldsymbol{a}_1)$，其中子向量$\boldsymbol{a}_0$和$\boldsymbol{a}_1$分别为$(p-1)\times 1$和$p\times 1$维列向量，线性组合

$$\sqrt{n}\boldsymbol{a}\ \ \widehat{\boldsymbol{M}}_i\ \ \widehat{\boldsymbol{W}}^{-1}\begin{bmatrix}\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^{a}\tilde{\varepsilon}_{i,t}\\ \vdots\\ \frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-r-1}^{a}\tilde{\varepsilon}_{i,t}\end{bmatrix} \tag{A4}$$

在 $n\to\infty$ 时服从渐近正态分布，且渐近方差为 $\boldsymbol{a}\ \ \boldsymbol{U}_i\boldsymbol{a}$ 。为让证明过程更简洁，可以将除去 $\sqrt{n}\boldsymbol{a}\ \ \widehat{\boldsymbol{M}}_i$ 的部分视为一个整体，并记作 $\boldsymbol{\Xi}_i$，另外有

$$\widetilde{\boldsymbol{\Xi}}_i\equiv\boldsymbol{W}^{-1}\cdot\begin{bmatrix}\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^{a}\tilde{\varepsilon}_{i,t}\\ \vdots\\ \frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-r-1}^{a}\tilde{\varepsilon}_{i,t}\end{bmatrix}=\begin{bmatrix}w_{1,1}^{-1}\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^{a}\tilde{\varepsilon}_{i,t}+\cdots+w_{1,r}^{-1}\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-r-1}^{a}\tilde{\varepsilon}_{i,t}\\ \vdots\\ w_{r,1}^{-1}\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^{a}\tilde{\varepsilon}_{i,t}+\cdots+w_{r,r}^{-1}\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-r-1}^{a}\tilde{\varepsilon}_{i,t}\end{bmatrix}.$$

考虑在 $r=1$ 时的情况，因为 $\widehat{\boldsymbol{W}}^{-1}=\widehat{\boldsymbol{W}}^{-1}-\boldsymbol{W}^{-1}+\boldsymbol{W}^{-1}$ 和 $\|\widehat{\boldsymbol{W}}^{-1}-\boldsymbol{W}^{-1}\|_F=O_p(p/\sqrt{n})$（Ma 等，2021），有

$$\frac{1}{n}\sum_{t=5}^{n}\tilde{Y}_{j,t}^{a}\widetilde{\boldsymbol{Y}}_{t-2}^{b}\boldsymbol{\Xi}_i^0=\frac{1}{n}\sum_{t=5}^{n}\tilde{Y}_{j,t}^{a}\widetilde{\boldsymbol{Y}}_{t-2}^{b}\widetilde{\boldsymbol{\Xi}}_i^0(1+o_p(1)),$$

其中，$\boldsymbol{\Xi}_i^0$ 和 $\widetilde{\boldsymbol{\Xi}}_i^0$ 分别为 $\boldsymbol{\Xi}_i$ 和 $\widetilde{\boldsymbol{\Xi}}_i$ 中所对应的子向量。另外，对于 $\widetilde{\boldsymbol{\Xi}}_i$ 而言，每一行都可视为权重不同的线性组合，故首先探讨 $r=1$ 时的情况，有

$$\begin{aligned}\frac{1}{n}\sum_{t=5}^{n}\tilde{Y}_{j,t}^{a}\widetilde{\boldsymbol{Y}}_{t-2}^{b}w_{1,1}^{-1}\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^{a}\tilde{\varepsilon}_{i,t}&=\left[\frac{1}{n}\sum_{t=5}^{n}(\boldsymbol{e}_j\ \widetilde{\boldsymbol{Y}}_t^{a}\widetilde{\boldsymbol{Y}}_{t-2}^{b}-\boldsymbol{e}_j\ \boldsymbol{\Sigma}_2^{a})\right]w_{1,1}^{-1}(\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^{a}\tilde{\varepsilon}_{i,t})\\&\quad+\frac{n-4}{n}\boldsymbol{e}_j\ \boldsymbol{\Sigma}_2^{a}w_{1,1}^{-1}(\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^{a}\tilde{\varepsilon}_{i,t})\\&\equiv B_1+B_2.\end{aligned}$$

对于 B_1，有

$$\mathrm{E}\left[\frac{1}{n}\sum_{t=5}^{n}(\boldsymbol{e}_j\ \widetilde{\boldsymbol{Y}}_t^{a}\widetilde{\boldsymbol{Y}}_{t-2}^{b}\boldsymbol{e}_k-\boldsymbol{e}_j\ \boldsymbol{\Sigma}_2^{a}\boldsymbol{e}_k)\right]^2\text{,,}\ O(\frac{1}{n}),$$

以及

$$B_1\text{,,}\ \|\frac{1}{n}\sum_{t=5}^{n}(\widetilde{\boldsymbol{Y}}_{t-2}^{b}\boldsymbol{e}_j\ \widetilde{\boldsymbol{Y}}_t^{a}-\boldsymbol{\Sigma}_2^{a}\ \ \boldsymbol{e}_j)\|_2\|\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^{a}\tilde{\varepsilon}_{i,t}\|_2=O_p(\frac{p}{n}). \tag{A5}$$

类似地，可知 $B_2=O_p(\frac{1}{\sqrt{n}})$ 和 $B_1=o_p(B_2)$ 。因此，在假设条件满足时，对于 $r..1$，关于线性组合式（A4）渐近正态性的证明可进一步转换为证明 $\mathrm{S}_{n,p}$ 服从渐近正态分布，其中，

$$S_{n,p}=\sqrt{n}\boldsymbol{a}\begin{bmatrix}\boldsymbol{I}_{p-1}\boldsymbol{\Sigma}_2^a\cdot\sum_{k=1}^{r}w_{1,k}^{-1}\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-k-1}^{a}\tilde{\varepsilon}_{i,t}+\cdots+\boldsymbol{I}_{p-1}\boldsymbol{\Sigma}_2^a\cdot\sum_{k=1}^{r}w_{r,k}^{-1}\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-k-1}^{a}\tilde{\varepsilon}_{i,t}\\ \boldsymbol{\Sigma}_1^b\cdot\sum_{k=1}^{r}w_{1,k}^{-1}\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-k-1}^{a}\tilde{\varepsilon}_{i,t}+\cdots+\boldsymbol{\Sigma}_1^b\cdot\sum_{k=1}^{r}w_{r,k}^{-1}\frac{1}{n}\sum_{t=2}^{n}\widetilde{\boldsymbol{Y}}_{t-k-1}^{a}\tilde{\varepsilon}_{i,t}\end{bmatrix}.$$

根据正文定义可知，$S_{n,p}=\boldsymbol{a}\ \boldsymbol{\Omega v}$，故有 $\mathrm{E}(S_{n,p})=0,\mathrm{Var}(S_{n,p})=\boldsymbol{a}\ \boldsymbol{\Omega\Sigma}_{\tilde{\boldsymbol{v}}^a\tilde{\boldsymbol{v}}^a}^{'}\boldsymbol{\Omega}\ \boldsymbol{a}$。因此，当 $n\to\infty$ 时有 $\mathrm{Var}(S_{n,p}/\sqrt{\boldsymbol{a}\ \boldsymbol{U}_i\boldsymbol{a}})\to 1$。由于 $S_{n,p}$ 的构成形式为有限数量的线性加权求和，若每个基本求和项都服从渐近正态分布，则 $S_{n,p}$ 自然也是服从渐近正态分布。对于 $S_{n,p}$ 的基本求和项而言，其渐近正态的证明过程与 Wang 等（2021）中的式（29）完全一致。至此，部分（*）已证明完毕。

关于（**）部分，有

$$\widehat{\boldsymbol{M}}_i\ \widehat{\boldsymbol{W}}^{-1}\widehat{\boldsymbol{M}}_i=\begin{bmatrix}\frac{1}{n}\sum_{t=5}^{n}\boldsymbol{I}_{p-1}\widetilde{\boldsymbol{Y}}_t^a\widetilde{\boldsymbol{Y}}_{t-2}^b & \cdots & \frac{1}{n}\sum_{t=r+4}^{n}\boldsymbol{I}_{p-1}\widetilde{\boldsymbol{Y}}_t^a\widetilde{\boldsymbol{Y}}_{t-r-1}^b\\ \frac{1}{n}\sum_{t=5}^{n}\widetilde{\boldsymbol{Y}}_{t-1}^b\widetilde{\boldsymbol{Y}}_{t-2}^b & \cdots & \frac{1}{n}\sum_{t=r+4}^{n}\widetilde{\boldsymbol{Y}}_{t-1}^b\widetilde{\boldsymbol{Y}}_{t-r-1}^b\end{bmatrix}\widehat{\boldsymbol{W}}^{-1}$$
$$\times\begin{bmatrix}\frac{1}{n}\sum_{t=5}^{n}\boldsymbol{I}_{p-1}\widetilde{\boldsymbol{Y}}_t^a\widetilde{\boldsymbol{Y}}_{t-2}^b & \cdots & \frac{1}{n}\sum_{t=r+4}^{n}\boldsymbol{I}_{p-1}\widetilde{\boldsymbol{Y}}_t^a\widetilde{\boldsymbol{Y}}_{t-r-1}^b\\ \frac{1}{n}\sum_{t=5}^{n}\widetilde{\boldsymbol{Y}}_{t-1}^b\widetilde{\boldsymbol{Y}}_{t-2}^b & \cdots & \frac{1}{n}\sum_{t=r+4}^{n}\widetilde{\boldsymbol{Y}}_{t-1}^b\widetilde{\boldsymbol{Y}}_{t-r-1}^b\end{bmatrix},$$

另外，对于前 $p-1$ 行的任一元素，即对于任意 $j_1,j_2=1,\cdots,p-1,$，有

$$\begin{aligned}\frac{1}{n}\sum_{t=5}^{n}\boldsymbol{e}_{j_1}\widetilde{\boldsymbol{Y}}_t^a\widetilde{\boldsymbol{Y}}_{t-2}^b\widehat{\boldsymbol{W}}^{-1}\frac{1}{n}\sum_{t=5}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^b\widetilde{\boldsymbol{Y}}_t^a\ \boldsymbol{e}_{j_2}&=\frac{1}{n}\sum_{t=5}^{n}\boldsymbol{e}_{j_1}\widetilde{\boldsymbol{Y}}_t^a\widetilde{\boldsymbol{Y}}_{t-2}^b(\widehat{\boldsymbol{W}}^{-1}-\boldsymbol{W}^{-1})\frac{1}{n}\sum_{t=5}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^b\widetilde{\boldsymbol{Y}}_t^a\ \boldsymbol{e}_{j_2}\\&\quad+\frac{1}{n}\sum_{t=5}^{n}\boldsymbol{e}_{j_1}\widetilde{\boldsymbol{Y}}_t^a\widetilde{\boldsymbol{Y}}_{t-2}^b\boldsymbol{W}^{-1}\frac{1}{n}\sum_{t=5}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^b\widetilde{\boldsymbol{Y}}_t^a\ \boldsymbol{e}_{j_2}\\&\equiv B_1^{'}+B_2^{'}.\end{aligned}$$

与前述 B_1 和 B_2 类似，易知 $B_1^{'}=o_p(B_2^{'})$。进一步可以构造出一个 $B_2^{'}$ 的上界，具体为

$$\begin{aligned}&(\frac{1}{n}\sum_{t=5}^{n}\boldsymbol{e}_{j_1}\widetilde{\boldsymbol{Y}}_t^a\widetilde{\boldsymbol{Y}}_{t-2}^b\boldsymbol{W}^{-1}-\boldsymbol{e}_{j_1}\boldsymbol{\Sigma}_2^a\boldsymbol{W}^{-1})(\frac{1}{n}\sum_{t=5}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^b\widetilde{\boldsymbol{Y}}_t^a\ \boldsymbol{e}_{j_2}-\boldsymbol{\Sigma}_2^a\ \boldsymbol{e}_{j_2})\\&+\boldsymbol{e}_{j1}\boldsymbol{\Sigma}_2^a\boldsymbol{W}^{-1}(\frac{1}{n}\sum_{t=5}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^b\widetilde{\boldsymbol{Y}}_t^a\ \boldsymbol{e}_{j_2}-\boldsymbol{\Sigma}_2^a\ \boldsymbol{e}_{j_2})\\&+(\frac{1}{n}\sum_{t=5}^{n}\boldsymbol{e}_{j_1}\widetilde{\boldsymbol{Y}}_t^a\widetilde{\boldsymbol{Y}}_{t-2}^b\boldsymbol{W}^{-1}-\boldsymbol{e}_{j_1}\boldsymbol{\Sigma}_2^a\boldsymbol{W}^{-1})\boldsymbol{\Sigma}_2^a\ \boldsymbol{e}_{j_2}\\&+\boldsymbol{e}_{j_1}\boldsymbol{\Sigma}_2^a\boldsymbol{W}^{-1}\boldsymbol{\Sigma}_2^a\ \boldsymbol{e}_{j_2}\\&\equiv Q_1+Q_2+Q_3+Q_4,\end{aligned}$$

与式（A5）类似，当 $p=o(\sqrt{n})$ 时，有

$$Q_1 = O_p(\frac{p}{n}), \qquad Q_2 = O_p(\frac{1}{\sqrt{n}}), \qquad Q_3 = O_p(\frac{1}{\sqrt{n}}),$$

进而，

$$\frac{1}{n}\sum_{t=5}^{n}\boldsymbol{e}_{j_1}\widetilde{\boldsymbol{Y}}_t^a\widetilde{\boldsymbol{Y}}_{t-2}^b\widehat{\boldsymbol{W}}^{-1}\frac{1}{n}\sum_{t=5}^{n}\widetilde{\boldsymbol{Y}}_{t-2}^b\widetilde{\boldsymbol{Y}}_t^a\ \boldsymbol{e}_{j_2} = Q_4 + o_p(1) = \boldsymbol{e}_{j_1}\boldsymbol{\Sigma}_2^a\boldsymbol{W}^{-1}\boldsymbol{\Sigma}_2^b\boldsymbol{e}_{j_2} + o_p(1).$$

相同地，对 $\widehat{\boldsymbol{M}}_i\ \widehat{\boldsymbol{W}}^{-1}\widehat{\boldsymbol{M}}_i$ 中的任一元素做以上运算，便可得到

$$\boldsymbol{K}_0(\widehat{\boldsymbol{M}}_i\ \widehat{\boldsymbol{W}}^{-1}\widehat{\boldsymbol{M}}_i)^{-1} \xrightarrow{p} \boldsymbol{I}_{2p-1}.$$

至此，定理2证明证毕。

删除[kylin]:

附表1　　情形一的500次数值模拟矩阵估计结果

p	n	r	SNR	范数误差		平均绝对误差	
				A	***B***	***A***	***B***
50	100	1	1.016	4.132 (0.100)	3.295 (0.071)	0.065 (0.002)	0.052 (0.001)
	200	1	1.182	3.717 (0.071)	2.984 (0.054)	0.059 (0.001)	0.047 (0.001)
	400	1	1.775	3.391 (0.051)	2.619 (0.054)	0.054 (0.001)	0.042 (0.001)
	1000	1	3.767	3.084 (0.024)	2.082 (0.045)	0.049 (0.000)	0.033 (0.001)
100	100	1	1.396	5.039 (0.062)	4.699 (0.027)	0.040 (0.000)	0.037 (0.000)
	200	1	3.130	4.634 (0.074)	4.042 (0.042)	0.037 (0.001)	0.032 (0.000)
	400	1	3.302	4.583 (0.050)	3.819 (0.030)	0.036 (0.000)	0.030 (0.000)
	1000	1	2.368	4.704 (0.027)	3.459 (0.034)	0.037 (0.000)	0.028 (0.000)
200	100	2	1.501	6.501 (0.005)	6.462 (0.005)	0.026 (0.000)	0.026 (0.000)
	200	1	2.223	6.912 (0.034)	6.072 (0.021)	0.028 (0.000)	0.024 (0.000)
	400	1	3.394	6.526 (0.058)	5.724 (0.023)	0.026 (0.000)	0.023 (0.000)
	1000	1	1.115	7.298 (0.035)	5.695 (0.027)	0.029 (0.000)	0.023 (0.000)
300	100	3	13.549	7.757 (0.011)	7.501 (0.010)	0.021 (0.000)	0.020 (0.000)
	200	7	1.605	7.833 (0.014)	7.896 (0.007)	0.021 (0.000)	0.021 (0.000)
	400	1	3.097	8.080 (0.050)	7.193 (0.025)	0.021 (0.000)	0.019 (0.000)
	1000	1	3.735	7.723 (0.040)	6.651 (0.020)	0.020 (0.000)	0.018 (0.000)

注：表中矩阵的范数误差指所对应矩阵估计误差的 F 范数，平均绝对误差是对应矩阵估计误差的 MAE，括号内为标准误。下同。

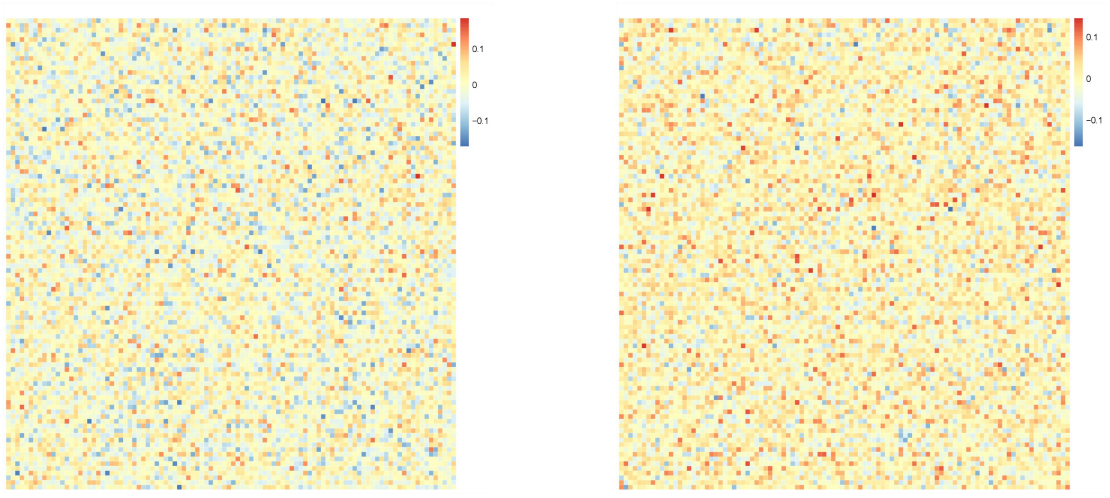
附表2　　情形二的500次数值模拟矩阵估计结果

p	n	r	SNR	范数误差		平均绝对误差	
				A	***B***	***A***	***B***
50	100	1	1.950	3.377 (0.105)	2.469 (0.103)	0.055 (0.001)	0.039 (0.002)
	200	1	2.169	3.296 (0.085)	2.359 (0.121)	0.054 (0.001)	0.038 (0.002)

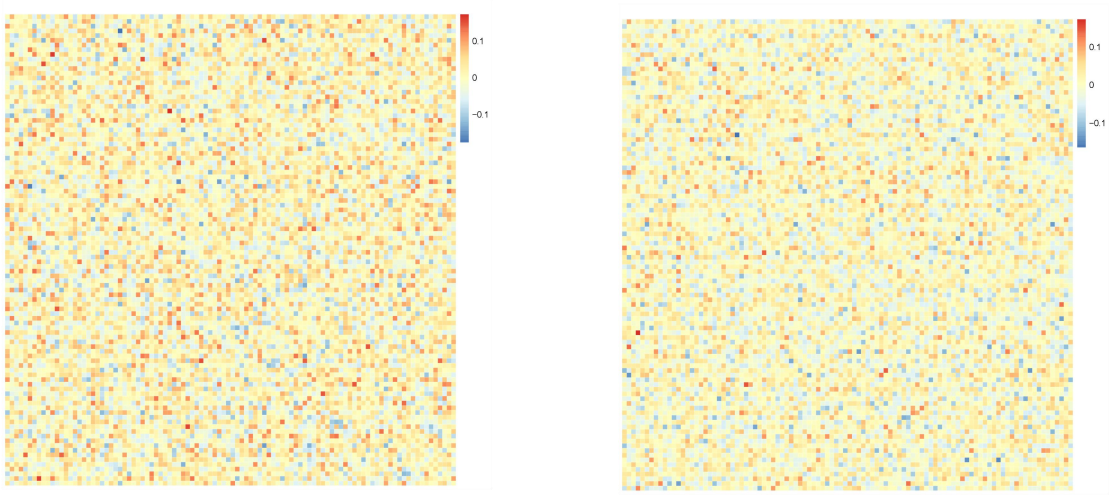
	400	1	1.962	3.328 (0.048)	2.379 (0.048)	0.055 (0.001)	0.038 (0.001)
	1000	1	3.053	3.077 (0.017)	1.788 (0.034)	0.052 (0.000)	0.028 (0.001)
100	100	1	1.973	4.760 (0.040)	3.520 (0.033)	0.039 (0.000)	0.028 (0.000)
	200	1	2.422	4.590 (0.080)	3.360 (0.102)	0.038 (0.000)	0.027 (0.001)
	400	1	2.668	4.484 (0.076)	3.160 (0.135)	0.037 (0.000)	0.025 (0.001)
	1000	1	2.386	4.509 (0.022)	3.175 (0.029)	0.037 (0.000)	0.025 (0.000)
200	100	1	7.135	6.216 (0.014)	4.279 (0.018)	0.026 (0.000)	0.017 (0.000)
	200	1	1.837	6.704 (0.028)	4.988 (0.027)	0.027 (0.001)	0.020 (0.001)
	400	1	2.526	6.377 (0.103)	5.147 (0.151)	0.025 (0.000)	0.021 (0.001)
	1000	1	1.188	7.159 (0.032)	6.024 (0.020)	0.028 (0.000)	0.024 (0.000)
300	100	3	6.618	7.692 (0.010)	4.740 (0.582)	0.022 (0.000)	0.013 (0.001)
	200	4	3.172	7.865 (0.020)	5.698 (0.016)	0.022 (0.000)	0.015 (0.000)
	400	1	2.168	7.996 (0.052)	5.966 (0.052)	0.022 (0.000)	0.016 (0.000)
	1000	1	1.024	9.564 (0.158)	7.858 (0.115)	0.026 (0.000)	0.021 (0.000)

删除[kylin]: **附录4**

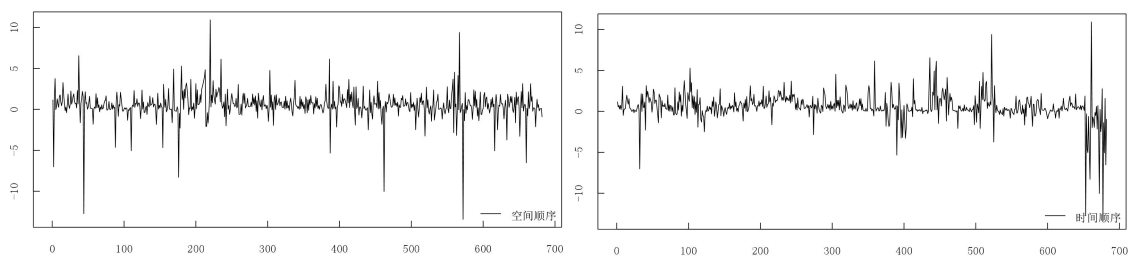
附录4展示数值模拟部分中两种情形下参数矩阵估计误差值的热力图（附图1和附图2），以及实证研究中产品市场差分序列值（附图3）和要素市场差分序列值（附图4）。



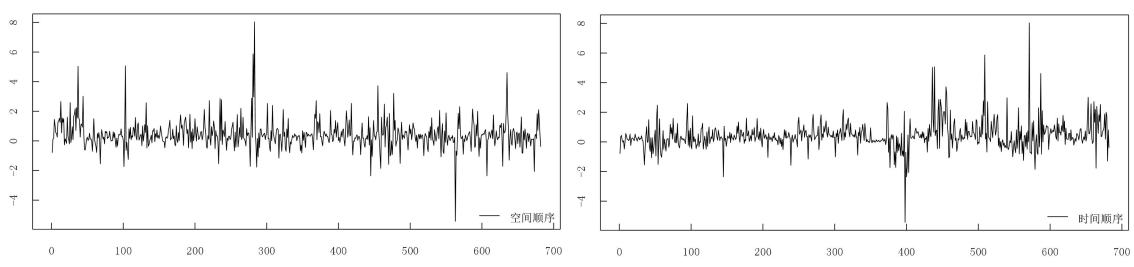
附图1　情形一参数矩阵 $\boldsymbol{A}$（左图）和 $\boldsymbol{B}$（右图）的估计误差热力图



附图2　情形二参数矩阵 ***A***（左图）和 ***B***（右图）的估计误差热力图



附图3　产品市场差分序列值



附图4　要素市场差分序列值

删除[kylin]:

参考文献

Ma Y, Guo S, Wang H. Sparse Spatio-temporal Autoregressions by Profiling and Bagging[J]. Journal of Econometrics, 2023, 232(1): 132-147.

Wang H, Luo X, Ling L. Semiparametric Spatio-temporal Models with Unknown and Banded Autoregressive Coefficient Matrices[J]. Mathematical Methods in the Applied Sciences, 2021.