arXiv:2304.06110v1 [stat.ME] 12 Apr 2023

基于小波的时变 STARMA 模型

Yangyang Chen[1], Pedro A. Morettin[1], *, Chang Chiann.[1]

[1]圣保罗大学统计系~

摘要

时空自回归移动平均（STARMA）模型常用于多元时间序列数据的研究，这类研究通常假设数据具有平稳性，但在实践中这一假设并非总能得到保证。一种可行的方法是考虑局部平稳过程。本文基于局部平稳假设，提出一种时变时空自回归移动平均（tvSTARMA）建模方法。我们将时变参数表示为小波基的线性组合，并提出了估计系数的程序。文中还展示了一些模拟结果，以及对美国中西部各州历史日降水记录的应用。

关键词 时空、局部平稳过程、时变、卡尔曼滤波器、小波

1 引言

自 Box 和 Jenkins（1970）提出季节性自回归积分移动平均（SARIMA）模型这一类模型，用于识别、估计和检验特定时间序列数据集的模型以来，该方法已成为分析具有常系数和方差的平稳及齐次非平稳时间序列的标准方法。拟合这些模型是为了更好地理解数据，或预测序列中的未来值（预测）。对于平稳的多元时间序列，通常采用向量自回归移动平均（VARMA）模型。

在时空地统计学背景下，经典模型是 STARMA（时空自回归移动平均）模型，它是 VARMA 模型的一个特例。Cliff 和 Ord（1975）以及 Martin 和 Oeppen（1975）最早采用了 STARMA 类模型；随后，Pfeifer 和 Deutsch（1980a、1980b、1981a、1981b、1981c）在 20 世纪 80 年代初提出了多种方法。然而，由于计算实现困难，该模型一度被人们遗忘。直到近些年，STARMA 模型才重新受到关注，并常用于地统计分析，以描述具有时空滞后线性依赖特征的空间局部化时间序列；也就是说，建模过程以在 N 个地理位置观测到的随机变量为特征，并且在每个位置上， T 都随时间进行观测。

*通讯作者。电子邮箱：pam@ime.usp.br（P. A. Morettin）。

在这种情况下，除了近期过去的数值影响更大之外，近处位置的影响也大于远处位置；这是通过指定空间权重矩阵实现的，该矩阵为最近邻赋予最高权重。STARMA 模型已广泛应用于不同类型的时空数据，例如房地产价格（Pace 等（2000））、交通流数据（Kamarianakis 和 Prastacos（2005））、损伤检测（Hu 等（2011））、区域银行存款（Kurt 和 Tunay（2015））、风能（Zou 等（2018））等。

与 ARIMA 类模型一样，STARMA模型也具有平稳性的假设，然而在实践中，这一假设并不一定成立。有许多众所周知的技术可以将非平稳时间序列转换为平稳时间序列，例如差分、对数变换、去趋势等。一个颇具吸引力的替代方案是局部平稳过程的思想，在这种过程中，该过程在较短的时间段内近似平稳，但其特征（协方差、参数等）会在整个时间段内逐渐变化。Dahlhaus（1996a、1996b、1996c、1997、2000）以及 Dahlhaus等人（1999）提出了许多估计量和渐近结果；关于局部平稳过程的综述可参见 Dalhaus（2012）。

基于Dahlhaus 提出的局部平稳过程的时变模型，多年来已被应用于许多不同的场景。例如，Chiann和Morettin（1999，2005）研究了线性系统中时变系数的估计；Sato等人（2007）提出了时变 VAR 模型的估计程序，并将其应用于功能性磁共振成像数据；Rohan和Ramanathan（2013）针对时变广义自回归条件异方差（GARCH）模型引入了非参数估计；Yousuf和 Ng（2021）提出了两种用于估计高维线性时变参数（TVP）模型的算法。本文基于系数的小波展开，提出一种时变时空自回归移动平均模型（tvSTARMA）。

本文内容安排如下。第 2 节介绍局部平稳过程和小波的基本背景。第 3 节和第 4 节提出时变时空自回归移动平均模型，并介绍相应的估计程序。第 5 节和第 6 节介绍一些模拟研究，第 7 节展示历史日降水记录的应用。第 8 节给出进一步的评论。

2 背景

2.1 局部平稳过程

如第 1 节所述，平稳性是时间序列分析的一项基本假设；然而，许多实际应用中的现象都呈现出非平稳行为。解决这一问题的一种方法是假设所涉及的过程是局部平稳的。

定义 2.1 (Dahlahus (2000)) 若高斯多元随机过程序列 $X_{t,\ T} = (X_{t\ T,}^{(1)},\ \ldots\ldots,\ X_{t\ T,}^{(d)})'(t = 1,\ \ldots\ldots,\ T)$ 存在如下表示，且其传递函数矩阵为 $\mathbf{A}^0$、均值函数向量为 μ，则称该序列为局部平稳过程

$$X_{t,T} = \mu(\frac{t}{T}) + \int_{-\pi}^{\pi} \exp(i\lambda t)\mathbf{A}^0_{t,T}(\lambda)d\xi(\lambda) \tag{1}$$

具有以下性质：

(i) $\xi(\lambda)$ 是定义在 $[-\pi, \pi]$ 上的复值高斯向量过程，且 $\xi_a(\lambda) = \xi_a(-\lambda)$, $\overline{E}[\xi_a(\lambda)] = 0$ 并且

$$E[d\xi_a(\lambda)d\xi_b(\lambda)] = \delta_{ab}\eta(\lambda + \mu)d\lambda d\mu, \tag{2}$$

其中，$\eta(\lambda) = \sum_{j=-\infty}^{\infty} \delta(\lambda + 2\pi j)$ 是狄拉克 δ 函数的周期 2π 延拓。

(ii) 存在一个常数 K 和一个 2π 周期矩阵值函数 $\mathbf{A}$:$[0,1]\times\mathbb{R}\to\mathbb{C}^{d\times d}$，满足 $\mathbf{A}(u,\lambda)=\mathbf{A}(u,-\lambda)$，且

$$\sup_{t,\lambda}\mid \mathbf{A}^{0}_{t,T}(\lambda)_{ab}-\mathbf{A}(t/T,\lambda)_{ab}\mid\leq KT^{-1} \tag{3}$$

对于所有 a, $b=1$, ..., d 和 $T\in\mathbb{N}$。$\mathbf{A}(u,\lambda)$ 和 $\mu(u)$ 假设在 μ 上连续。

函数 $f(u,\lambda):=\mathbf{A}(u,\lambda)\mathbf{A}(u,-\lambda)'$ 是该过程的时变谱密度矩阵。

如果 $X_{t,T}$ 是一个时变向量自回归移动平均模型，则它由差分方程定义

$$\sum_{j=0}^{p}\Phi_j(\frac{t}{T})\left[X_{t-j,T}-\mu(\frac{t-j}{T})\right]=\sum_{j=0}^{q}\Psi_j(\frac{t}{T})\epsilon_{t-j}, \tag{4}$$

其中 ϵ_t 相互独立且同分布，均值为零，协方差矩阵为 $\mathbf{I}_d$，且 $\Phi_0(u)\equiv\Psi_0(u)\equiv\mathbf{I}_d$。对于 $z\in\mathbb{C}$，设 $\Phi(u,z)=\sum_{j=0}^{p}\Phi_j(u)z^j$ 且 $\Psi(u,z)=\sum_{j=0}^{q}\Psi_j(u)z^j$。若 $det(\Phi(u,z))\neq 0$ 对所有 $|z|\leq 1+c$ 都成立，且 $c>0$ 在 u 上一致，并且 $\Phi_j(u)$ 和 $\Psi_j(u)$ 的所有元素均在 u 上连续，则该差分方程的解具有无限时变 MA 表示，也就是说，该解是形式如 (1) 的局部平稳过程。该过程的时变谱密度为

$$f(u,\lambda)=\frac{1}{2\pi}\frac{\Psi(u,e^{i\lambda})\Psi(u,e^{-i\lambda})'}{\Phi(u,e^{i\lambda})\Phi(u,e^{-i\lambda})'}. \tag{5}$$

由于 STARMA 模型是 VARMA 模型的特例，因此将要提出的时变 STARMA 模型是一个局部平稳过程；为简便起见，以下假设 $\mu(\cdot)=0$。

2.2 小波

小波分析与傅里叶分析密切相关，因为二者都能将给定函数表示为基函数之和。小波基的主要特征是，函数 $f\in L^2(\mathbb{R})$ 可以通过对尺度函数 $\phi(t)$ 和 / 或母小波 $\psi(t)$ 进行二进制伸缩 2^j 以及二进制平移 $k2^{-j}$ 来生成。因此，小波基由函数 $\{\phi_{j,k}(t)\cup\psi_{j,k}(t)\ j、\ k\in\mathbb{Z}\}\in L^2(\mathbb{R})$ 组成，其中

$$\phi_{j,k}(t)=2^{j/2}\phi(2^jt-k), \tag{6}$$

and

$$\psi_{j,k}(t)=2^{j/2}\psi(2^jt-k). \tag{7}$$

获得尺度函数的一种方法是求方程的解

$$\phi(t)=\sqrt{2}\sum_k l_k\phi(2t-k), \tag{8}$$

而 $\psi(t)$ 由 $\phi(t)$ 通过下式得到

$$\psi(t)=\sqrt{2}\sum_k h_k\phi(2t-k), \tag{9}$$

事实上，l_k 和 h_k 分别是低通滤波器和高通滤波器的系数，由下式给出

$$l_k=\sqrt{2}\int_{-\infty}^{\infty}\phi(t)\phi(2t-k)dt, \tag{10}$$

$$h_k=\sqrt{2}\int_{-\infty}^{\infty}\psi(t)\phi(2t-k)dt. \tag{11}$$

本文将使用墨西哥帽小波，其母函数定义为

$$\psi(t)=\frac{2}{\sqrt{3}\pi^{1/4}}(1-t^2)e^{-t^2/2}. \tag{12}$$

多分辨率分析（MRA）使我们能够在不同的分辨率层级上分析数据。粗分辨率的数据包含低频分量的信息，并保留原始信号的主要特征；而细分辨率的数据则保留高频分量的信息。

设 V_j 和 W_j 分别是由 $\{\phi_{j,k}, k = 0, ..., 2^j - 1\}$ 和 $\{\psi_{j,k}, k = 0, ..., 2^j - 1\}$ 生成的闭子空间，则多分辨率分析具有以下性质：

(i) $... \subset V_{-1} \subset V_0 \subset V_1 \subset ...$
(ii) $L^2(\mathbb{R}) = \bigcup_j V_j$ (iii) $\bigcap_j V_j = \{0\}$ (iv) $x(t) \in V_j \Leftrightarrow x(2t) \in V_{j+1}$, $\forall j$ (v) $V_{j+1} = V_j \oplus W_j$, $W_j \perp V_j$.

上述性质意味着 $W_j = V_{j+1} \ominus V_j$，因此我们得到

$$L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} W_j = V_0 \oplus \bigoplus_{j\geq 0} W_j = V_{J_0} \oplus \bigoplus_{j\geq J_0} W_j \tag{13}$$

对于某个 J_0 整数，任意 $f(t) \in L^2(\mathbb{R})$ 均可表示为

$$f(t) = \sum_k c_{J_0,k}\phi_{J_0,k}(t) + \sum_{j\geq J_0}\sum_k d_{j,k}\psi_{j,k}(t), \tag{14}$$

对于某个较粗的尺度 J_0 （通常取为零），其中

$$c_{j,k} = \int_{-\infty}^{\infty} f(t)\phi_{j,k}(t)dt,$$
$$d_{j,k} = \int_{-\infty}^{\infty} f(t)\psi_{j,k}(t)dt.$$

3 时变 STARMA 模型

3.1 空间权重矩阵

在构建 STARMA 模型时，需要定义空间滞后算子。设 $Z_i(t)$ 是位置 i 在时间 t 和 $L^{(l)}$处的一个随机变量，则空间阶数为 l 的空间滞后算子定义为

$$L^{(l)}Z_i(t) = \begin{cases} Z_i(t), & \text{if } l = 0, \\ \sum_{j=1}^n w_{ij}^{(l)} Z_j(t), & \text{if } l > 0, \end{cases} \tag{15}$$

其中 $w_{ij}^{(l)}$ 是阶数为 l 的权重，满足

$$\sum_{j=1}^n \omega_{ij}^{(l)} = 1. \tag{16}$$

矩阵的每个元素都反映两个区域之间的空间关系： i 和 j 以及 $w_{ij}^{(l)} = 0$ 在 $i = j$时的空间关系；也就是说，矩阵的主对角线为零，其余元素均为正数。用矩阵形式表示，设 $\mathbf{Z}(t) = [Z_1(t)..., Z_n(t)]'$为由 n 个观测值组成的向量，

$$L^{(l)}\mathbf{Z}(t) = \begin{cases} W^{(0)}\mathbf{Z}(t) = \mathbf{I}_n\mathbf{Z}(t), & \text{if } l = 0, \\ W^{(l)}\mathbf{Z}(t), & \text{if } l > 0, \end{cases} \tag{17}$$

其中， $W^{(l)}$ 是一个由权重 $w_{ij}^{(l)}$ 组成的 $n \times n$ 方矩阵，使得每行中各项之和为一。

3.2 基于小波的时变 STAR 模型

经典时空自回归移动平均（$p_{\lambda_1,...,\lambda_p}, q_{m_1,...,^{m_q}}$）模型可定义为

$$Z_i(t) = \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\phi_{sl}L^{(l)}Z_i(t-s) - \sum_{s=1}^{q}\sum_{l=0}^{m_s}\theta_{sl}L^{(l)}\epsilon_i(t-s) + \epsilon_i(t),\ i=1,...,n,\ t=1,...,T, \tag{18}$$

或写成矩阵形式：

$$\mathbf{Z}(t) = \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\phi_{sl}W^{(l)}\mathbf{Z}(t-s) - \sum_{s=1}^{q}\sum_{l=0}^{m_s}\theta_{sl}W^{(l)}\boldsymbol{\epsilon}(t-s) + \boldsymbol{\epsilon}(t), \tag{19}$$

其中， p 为自回归阶数， q 为移动平均阶数， λ_s 为第 s 个自回归项的空间阶数， m_s 为第 s 个移动平均项的空间阶数。参数 ϕ_{sl} 和 θ_{sl} 分别表示自回归参数和移动平均参数在时间滞后 s 和空间滞后 l 下的参数，且 $\boldsymbol{\epsilon}(t)$为相互独立且服从正态分布的随机误差，其 $E[\boldsymbol{\epsilon}(t)] = \mathbf{0}$ ，协方差矩阵为$\mathbf{I}_N\sigma^2$。

设$\mathbf{Z}(\frac{t}{T}) = [Z_1(\frac{t}{T}),...,Z_n(\frac{t}{T})]'$为一个 n 维时间序列，包含 T 个观测值。时变 STAR（tvS-TAR）模型定义为

$$\mathbf{Z}\left(\frac{t}{T}\right) = \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\phi_{sl}\left(\frac{t}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{t-s}{T}\right) + \boldsymbol{\epsilon}\left(\frac{t}{T}\right),\ t=1,...,T, \tag{20}$$

其中， $\boldsymbol{\epsilon}(\frac{t}{T})$为均值为零的独立同分布高斯向量， $\phi_{sl}(\frac{t}{T})$为时间滞后 s 和空间滞后 l 下的时变参数， $W^{(l)}$ 为空间阶数为 l 的空间权重矩阵。

尽管小波基的形状不规则，但仍能完美重构具有线性及更高阶多项式形状的函数。其思路是将时变参数 $\phi_{sl}(\frac{t}{T})$ 展开为小波展开，即

$$\phi_{sl}\left(\frac{t}{T}\right) = \sum_{j=-1}^{\infty}\sum_{k=0}^{2^j-1}\beta_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right). \tag{21}$$

注意，为简化起见，我们定义 $\psi_{-1,0}(\frac{t}{T}) = \phi_{0,0}(\frac{t}{T})$，即 $j,k=0$ 的尺度函数。

然后，基于小波的时变时空自回归模型表示为

$$\begin{aligned}
\mathbf{Z}\left(\frac{t}{T}\right) &= \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\sum_{j=-1}^{\infty}\sum_{k=0}^{2^j-1}\beta_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{t-s}{T}\right) + \boldsymbol{\epsilon}\left(\frac{t}{T}\right) \\
&= \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\sum_{j=-1}^{J-1}\sum_{k=0}^{2^j-1}\beta_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{t-s}{T}\right) \\
&+ \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\sum_{j\geq J}^{\infty}\sum_{k=0}^{2^j-1}\beta_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{t-s}{T}\right) + \boldsymbol{\epsilon}\left(\frac{t}{T}\right) \\
&= \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\sum_{j=-1}^{J-1}\sum_{k=0}^{2^j-1}\beta_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{t-s}{T}\right) + \boldsymbol{\nu}\left(\frac{t}{T}\right),
\end{aligned} \tag{22}$$

其中

$$\boldsymbol{\nu}\left(\frac{t}{T}\right) = \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\sum_{j\geq J}^{\infty}\sum_{k=0}^{2^j-1}\beta_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{t-s}{T}\right) + \boldsymbol{\epsilon}\left(\frac{t}{T}\right) \tag{23}$$

且 $W^{(l)}$ 是一个权重矩阵，每行的和均为一。注意， $J-1$ 是满足 $2^{J-1} \leq \sqrt{T} \leq 2^J$ 的最精细分辨率层级，因此，$\sqrt{n}2^J/2 \leq \sqrt{n}T \leq \sqrt{n}2^J$ 满足 $n \geq 2$。

（22）的线性形式可写为

$$\mathbf{Z} = \boldsymbol{\Psi}\boldsymbol{\beta} + \boldsymbol{\nu}, \tag{24}$$

其中

$$\mathbf{Z} = \left[\text{vec}\left(\mathbf{Z}\left(\frac{p+1}{T}\right)\right)', ..., \text{vec}\left(\mathbf{Z}\left(\frac{T}{T}\right)\right)'\right]', \tag{25}$$

$$\boldsymbol{\nu} = \left[\text{vec}\left(\boldsymbol{\nu}\left(\frac{p+1}{T}\right)\right)', ..., \text{vec}\left(\boldsymbol{\nu}\left(\frac{T}{T}\right)\right)'\right]', \tag{26}$$

$$\boldsymbol{\Psi} = \left[\Psi_{-1,0}, \Psi_{0,0}, ..., \Psi_{J-1,2^{J-1}-1}\right] \tag{27}$$

and

$$\Psi_{j,k} = \begin{bmatrix} \psi_{j,k}\left(\frac{p+1}{T}\right)\sum_{m=1}^{n} w_{1,m}^{(0)} Z_m\left(\frac{p}{T}\right) & \cdots & \psi_{j,k}\left(\frac{p+1}{T}\right)\sum_{m=1}^{n} w_{1,m}^{(\lambda_p)} Z_m\left(\frac{1}{T}\right) \\ \vdots & & \vdots \\ \psi_{j,k}\left(\frac{T}{T}\right)\sum_{m=1}^{n} w_{1,m}^{(0)} Z_m\left(\frac{T-1}{T}\right) & \cdots & \psi_{j,k}\left(\frac{T}{T}\right)\sum_{m=1}^{n} w_{1,m}^{(\lambda_p)} Z_m\left(\frac{T-p}{T}\right) \\ \psi_{j,k}\left(\frac{p+1}{T}\right)\sum_{m=1}^{n} w_{2,m}^{(0)} Z_m\left(\frac{p}{T}\right) & \cdots & \psi_{j,k}\left(\frac{p+1}{T}\right)\sum_{m=1}^{n} w_{2,m}^{(\lambda_p)} Z_m\left(\frac{1}{T}\right) \\ \vdots & & \vdots \\ \psi_{j,k}\left(\frac{T}{T}\right)\sum_{m=1}^{n} w_{n,m}^{(0)} Z_m\left(\frac{T-1}{T}\right) & \cdots & \psi_{j,k}\left(\frac{T}{T}\right)\sum_{m=1}^{n} w_{n,m}^{(\lambda_p)} Z_m\left(\frac{T-p}{T}\right) \end{bmatrix}, \tag{28}$$

对于 $j = -1, ..., J-1$, $k = 0, ..., 2^j - 1$, 其中 $\boldsymbol{\beta}$ 是包含小波系数的 $s_a(2^J)$ 向量，且 $s_a = \sum_{s=1}^{p}(1+\lambda_s)$。

3.3 估计

参数的数量取决于空间阶数 λ_s（ $s = 1,..,p$）以及分辨率级别 J。对于规模非常大的数据集，需要估计的参数很多，此时最小二乘法是一种计算效率较高的估计方法。

假设矩阵 $\boldsymbol{\Psi}'\boldsymbol{\Psi}$ 是正定的，则 $\boldsymbol{\beta}$ 的最小二乘估计量为

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}'\mathbf{Z}. \tag{29}$$

我们考虑以下假设：

A.1 (Dahlhaus 等人 (1999)) 函数 ϕ_{kl} 为实值、有界，并属于以下函数集合：

$$\mathscr{F} = \left\{ f(x) : f(x) = \sum_{j=-1}^{\infty}\sum_{k=0}^{\infty} \beta_{j,k}\psi_{j,k}(x) \,\middle|\, \|\beta_{..}\|_{d,p,q} < \infty \right\}, \tag{30}$$

其中

$$\|\beta_{..}\|_{d,p,q} = \left[\sum_{j\geq -1}\left(2^{jup}\sum_{k=0}^{2^j-1}|\beta_{j,k}|^p\right)^{q/p}\right]^{1/q}, \tag{31}$$

$u = d + 1/2 - 1/\tilde{p} > 1$，其中 $\tilde{p} = \min\{p, 2\}$。注意， d 表示光滑度， $1 < p$， $q \leq \infty$ 指定衡量光滑度所用的范数。

A.2 函数 $\phi(t)$ 和 $\psi(t)$ 属于 $C^r[0,1]$，正则度为 $r > d$，且具有紧支撑。此外， $\int \phi(t)dt = 1$，且 $\int \psi(t)^k dt = 0$，对 $0 \leq k \leq r$ 成立。

A.3 存在某个 $\gamma \geq 0$，满足 $|\text{cum}_n(\nu_t)| \leq A^n(n!)^{1+\gamma}$，对所有 n， t 均成立，其中 A 为正数，且 $\nu_t = \nu(\frac{t}{T})$。这里，$\text{cum}_n(\nu_t)$ 表示 ν_t 的 n 阶累积量

A.4 $\mathbf{Z}(\frac{t}{T})$ 是局部平稳过程。

命题 1 假设 A.1-A.4 成立，则

（i.） $E(\hat{\boldsymbol{\beta}}) = O((nT)^{-1/2})$。（ii.）
$E[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'] = O((nT)^{-1})$。

证明。 请参见附录。

命题 2 假设 A.1-A.4 成立，则有

$$\sqrt{N}\mathbf{H}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{D} \mathcal{N}_K(\mathbf{0}, \boldsymbol{\Gamma}), \tag{32}$$

其中 D 表示依分布收敛和协方差矩阵

$$\boldsymbol{\Gamma} = \lim_{N\to\infty} N\mathbf{H}E[\boldsymbol{\Psi}'\boldsymbol{\Psi}]^{-1}\mathbf{H}', \tag{33}$$

其中， $N = n(T-p)$ 和 $\mathbf{H}$ 是一个有 K 行的矩阵。

证明。 参见附录。

4 基于小波的时变 STARMA 模型

4.1 模型

设$\mathbf{Z}\left(\frac{t}{T}\right) = [Z_1(\frac{t}{T})\ Z_2(\frac{t}{T})..., Z_n(\frac{t}{T})]'$为一个 n维时变时空自回归移动平均模型（tvSTARMA），且为多元局部平稳模型，其定义为

$$\mathbf{Z}\left(\frac{t}{T}\right) = \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\phi_{sl}\left(\frac{t}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{t-s}{T}\right) - \sum_{s=1}^{q}\sum_{l=0}^{m_s}\theta_{sl}\left(\frac{t}{T}\right)W^{(l)}\boldsymbol{\epsilon}\left(\frac{t-s}{T}\right) + \boldsymbol{\epsilon}\left(\frac{t}{T}\right),\ t = 1, ..., T, \tag{34}$$

其中， $\boldsymbol{\epsilon}(\frac{t}{T})$ 是均值为零的独立同分布高斯向量， $\phi_{sl}(\frac{t}{T})$ 和 $\theta_{sl}(\frac{t}{T})$ 分别是时间滞后 s 和空间滞后 l 处的时变参数，$W^{(l)}$ 是阶数为 l的空间权重矩阵。

时变参数 $\phi_{sl}(\frac{t}{T})$ 和 $\theta_{sl}(\frac{t}{T})$ 可表示为小波展开：

$$\phi_{sl}\left(\frac{t}{T}\right) = \sum_{j=-1}^{\infty}\sum_{k=0}^{2^j-1} a_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right), \tag{35}$$

$$\theta_{sl}\left(\frac{t}{T}\right) = \sum_{j=-1}^{\infty}\sum_{k=0}^{2^j-1} b_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right). \tag{36}$$

将方程（35）、（36）代入（34），可得基于小波的 tvSTARMA 模型：

$$\begin{aligned}
\mathbf{Z}\left(\frac{t}{T}\right) &= \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\sum_{j=-1}^{\infty}\sum_{k=0}^{2^j-1} a_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{t-s}{T}\right) \\
&- \sum_{s=1}^{q}\sum_{l=0}^{m_s}\sum_{j=-1}^{\infty}\sum_{k=0}^{2^j-1} b_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\boldsymbol{\epsilon}\left(\frac{t-s}{T}\right) + \boldsymbol{\epsilon}\left(\frac{t}{T}\right) \\
&= \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\sum_{j=-1}^{J-1}\sum_{k=0}^{2^j-1} a_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{t-s}{T}\right) \\
&- \sum_{s=1}^{q}\sum_{l=0}^{m_s}\sum_{j=-1}^{J-1}\sum_{k=0}^{2^j-1} b_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\boldsymbol{\epsilon}\left(\frac{t-s}{T}\right)
\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\sum_{j\geq J}^{\infty}\sum_{k=0}^{2^j-1} a_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{t-s}{T}\right)\\
&- \sum_{s=1}^{q}\sum_{l=0}^{m_s}\sum_{j\geq J}^{\infty}\sum_{k=0}^{2^j-1} b_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\boldsymbol{\epsilon}\left(\frac{t-s}{T}\right)+\boldsymbol{\epsilon}\left(\frac{t}{T}\right)\\
&= \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\sum_{j=-1}^{J-1}\sum_{k=0}^{2^j-1} a_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{t-s}{T}\right)\\
&- \sum_{s=1}^{q}\sum_{l=0}^{m_s}\sum_{j=-1}^{J-1}\sum_{k=0}^{2^j-1} b_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\boldsymbol{\epsilon}\left(\frac{t-s}{T}\right)+\boldsymbol{\nu}\left(\frac{t}{T}\right), \qquad (37)
\end{aligned}
$$

其中

$$
\begin{aligned}
\boldsymbol{\nu}\left(\frac{t}{T}\right) &= \sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}\sum_{j\geq J}^{\infty}\sum_{k=0}^{2^j-1} a_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{t-s}{T}\right)\\
&- \sum_{s=1}^{q}\sum_{l=0}^{m_s}\sum_{j\geq J}^{\infty}\sum_{k=0}^{2^j-1} b_{j,k}^{sl}\psi_{j,k}\left(\frac{t}{T}\right)W^{(l)}\boldsymbol{\epsilon}\left(\frac{t-s}{T}\right)+\boldsymbol{\epsilon}\left(\frac{t}{T}\right). \qquad (38)
\end{aligned}
$$

4.2 估计

与 STARMA 模型的情况一样，tvSTARMA 模型的估计可以采用线性估计量和非线性估计量。最大似然估计是一种广泛使用的技术，但当待估参数数量较多时，计算开销很大；而且对初始值的选择较为敏感，因此优化算法可能会收敛到局部最小值，甚至无法收敛。

另一种方法是卡尔曼滤波器（Kalman（1960年）），它与以下状态空间模型相关:

$$
\begin{aligned}
\mathbf{x}_{t+1} &= F_t\mathbf{x}_t+\Gamma_t\mathbf{w}_{t+1}, \qquad (39)\\
\mathbf{z}_t &= M_t\mathbf{x}_t+\mathbf{v}_t, \qquad (40)
\end{aligned}
$$

其中，(39) 是系统的状态方程，(40) 是系统的观测方程，$\mathbf{x}_t$ 是时刻 t 的状态变量，$\mathbf{z}_t$ 是 t 时刻的观测值，F_t，Γ_t 以及 M_t 分别是状态转移矩阵、控制输入矩阵和观测矩阵。向量$\mathbf{v}_t$ 和 $\mathbf{w}_t$ 是被假设为服从高斯分布的误差。该算法提供递归估计公式，用于在最小化平方误差的意义下估计过程的状态。

采用卡尔曼滤波器对 tvSTARMA 模型进行自适应参数估计时，首先将 (37) 重写为线性形式:

$$
\mathbf{Z}\left(\frac{t}{T}\right)=\mathbf{Y}\left(\frac{t}{T}\right)\mathbf{c}+\boldsymbol{\nu}\left(\frac{t}{T}\right), \qquad (41)
$$

其中，$\boldsymbol{\nu}\left(\frac{t}{T}\right)$ 假设服从均值为零且协方差矩阵为 Σ的多元正态分布，

$$
\begin{aligned}
s_a &= \sum_{s=1}^{p}(1+\lambda_s),\\
s_m &= \sum_{s=1}^{q}(1+m_s),\\
\mathbf{c}_{2^J[s_a+s_m]\times 1} &= \begin{bmatrix}\mathbf{a}\\ \mathbf{b}\end{bmatrix},\\
\mathbf{a}'_{2^J[s_a]\times 1} &= [a_{-1,0}^{10}\ a_{0,0}^{10}\ \ldots\ a_{J-1,2^{J-1}-1}^{10}\ \ldots\ a_{J-1,2^{J-1}-1}^{p\lambda_p}],
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{Y}\left(\underline{\begin{matrix} t \\ T \end{matrix}}\right) \quad \mathbf{b}'_{2^J\,[s_m]\times 1} &= \left[b^{10}_{-1,0}\ b^{10}_{0,0}\ \ldots\ b^{10}_{J-1,2^{J-1}-1}\ \ldots\ b^{qm_q}_{J-1,2^{J-1}-1}\right], \\
&= \left[\mathbf{diag}_n \otimes \mathbf{1}_{1\times[s_a+s_m]}\mathbf{WD}\left(\underline{\begin{matrix} t \\ T \end{matrix}}\right)\right] \otimes \Psi\left(\underline{\begin{matrix} t \\ T \end{matrix}}\right), \text{其中,} \\
{\scriptstyle n\times 2^J[s_a+s_m]} &= \left[\psi_{-1,0}\left(\underline{\begin{matrix} t \\ T \end{matrix}}\right)\ \psi_{0,0}\left(\underline{\begin{matrix} t \\ T \end{matrix}}\right)\ \ldots\ \psi_{J-1,2^{J-1}-1}\left(\underline{\begin{matrix} t \\ T \end{matrix}}\right)\right], \\
\Psi\left(\frac{t}{T}\right)'_{2^J\times 1}
\end{aligned}
$$

$$
\mathbf{W}_{n[s_a+s_m]\times n[s_a+s_m]} = \begin{bmatrix} \mathbf{W}^{ar} & \\ & \mathbf{W}^{ma} \end{bmatrix},
$$

$$
\mathbf{W}^{ar}_{n\times n[s_a]} = \begin{bmatrix}
W_{(0)} & & & & & & & & \\
& \ddots & & & & & & & \\
& & W_{(\lambda_1)} & & & & & & \\
& & & W_{(0)} & & & & & \\
& & & & \ldots & & & & \\
& & & & & W_{(\lambda_2)} & & & \\
& & & & & & \ddots & & \\
& & & & & & & W_{(\lambda_p)}
\end{bmatrix},
$$

$$
\mathbf{W}^{ma}_{n\times n[s_m]} = \begin{bmatrix}
W_{(0)} & & & & & & & & \\
& \ddots & & & & & & & \\
& & W_{(m_1)} & & & & & & \\
& & & W_{(0)} & & & & & \\
& & & & \ldots & & & & \\
& & & & & W_{(m_2)} & & & \\
& & & & & & \ddots & & \\
& & & & & & & W_{(m_q)}
\end{bmatrix},
$$

$$
\mathbf{D}\left(\underline{\begin{matrix} t \\ T \end{matrix}}\right)_{n[s_a+s_m]\times[s_a+s_m]} = \begin{bmatrix} \mathbf{X}\left(\underline{\begin{matrix} t \\ T \end{matrix}}\right) & \\ & \mathbf{E}\left(\underline{\begin{matrix} t \\ T \end{matrix}}\right) \end{bmatrix},
$$

$$
\mathbf{X}\left(\frac{t}{T}\right)_{n[s_a]\times s_a} = \begin{bmatrix}
\mathbf{diag}_{(1+\lambda_1)} \otimes \mathbf{Z}(\frac{t-1}{T}) & & & \\
& \mathbf{diag}_{(1+\lambda_2)} \otimes \mathbf{Z}(\frac{t-2}{T}) & & \\
& & \ddots & \\
& & & \mathbf{diag}_{(1+\lambda_p)} \otimes \mathbf{Z}(\frac{t-p}{T})
\end{bmatrix},
$$

$$
\mathbf{E}\left(\frac{t}{T}\right)_{n[s_m]\times s_m} = \begin{bmatrix}
\mathbf{diag}_{(1+m_1)} \otimes \boldsymbol{\epsilon}(\frac{t-1}{T}) & & & \\
& \mathbf{diag}_{(1+m_2)} \otimes \boldsymbol{\epsilon}(\frac{t-2}{T}) & & \\
& & \ddots & \\
& & & \mathbf{diag}_{(1+m_q)} \otimes \boldsymbol{\epsilon}(\frac{t-q}{T})
\end{bmatrix}.
$$

根据 Cipra 和 Motykova（1987），用于估计参数 $\mathbf{c}$ 的状态空间表示（39）和（40）可以写为

$$
\mathbf{c}_{t+1} \quad = \quad \mathbf{c}_t, \tag{42}
$$

$$\mathbf{z}_t \quad = \quad \mathbf{y}_t \mathbf{c}_t + \boldsymbol{\epsilon}_t, \tag{43}$$

其中 $F_t = \mathbf{I}_{2^J[s_a+s_m]}$, $\Gamma_t = \mathbf{0}_{2^J[s_a+s_m]}$, $\mathbf{z}_t = \mathbf{Z}(\frac{t}{T})$, $M_t = \mathbf{y}_t = \mathbf{Y}(\frac{t}{T})$, $\mathbf{v}_t = \boldsymbol{\nu}_t = \boldsymbol{\nu}(\frac{t}{T})$。然后，递归方程具有以下形式：

$$\begin{aligned}
\hat{\mathbf{c}}_{t+1} &= \hat{\mathbf{c}}_t + \mathbf{P}_{t+1}\mathbf{y}'_{t+1}\hat{\boldsymbol{\Sigma}}_t^{-1}(\mathbf{z}_{t+1} - \mathbf{y}_{t+1}\hat{\mathbf{c}}_t), && (44)\\
\mathbf{P}_{t+1} &= \mathbf{P}_t - \mathbf{P}_t\mathbf{y}'_{t+1}(\mathbf{y}_{t+1}\mathbf{P}_t\mathbf{y}'_{t+1} + \hat{\boldsymbol{\Sigma}}_t)^{-1}\mathbf{y}'_{t+1}\mathbf{P}_t, && (45)\\
\hat{\boldsymbol{\Sigma}}_{t+1} &= \frac{1}{t+1-(s_a+s_m)}\{[t-(s_a+s_m)]\hat{\boldsymbol{\Sigma}}_t + (\mathbf{z}_{t+1} - \mathbf{y}_{t+1}\hat{\mathbf{c}}_{t+1})(\mathbf{z}_{t+1} - \mathbf{y}_{t+1}\hat{\mathbf{c}}_{t+1})'\}, && (46)\\
\hat{\boldsymbol{\nu}}_{t+1} &= \mathbf{z}_{t+1} - \mathbf{y}_{t+1}\hat{\mathbf{c}}_{t+1}. && (47)
\end{aligned}$$

注意，如果没有关于参数的先验信息，则时间 t_0 处估计值的初始值可以选为

$$\hat{\mathbf{c}}_{t_0} = \mathbf{0}_{2^J[s_a+s_m]\times 1},\ \mathbf{P}_{t_0} = \mathbf{I}_{2^J[s_a+s_m]\times 1},\ \hat{\boldsymbol{\Sigma}}_{t_0} = h\mathbf{I}_{n\times n}, \tag{48}$$

其中， h 是一个较小的正数。

5 模拟

本节给出一些模拟示例，以评估所提出的估计程序的性能。

5.1 仿真过程

模拟包括以下步骤：

[1.]令 $n = 15$ 个样本位置（如图1所示），然后模拟 $M = 1000$ 组长度为 $T = 1024$的 n 时间序列。

[2.] 每个样本数据 $\mathbf{z} = [\mathbf{z}(\frac{1}{T}), ..., \mathbf{z}(\frac{T}{T})]$ 均由一个 tvSTAR(1_1) 过程模拟生成

$$\mathbf{z}\left(\frac{t}{T}\right) = \phi_{10}(t)W^{(0)}\mathbf{z}\left(\frac{t-1}{T}\right) + \phi_{11}(t)W^{(1)}\mathbf{z}\left(\frac{t-1}{T}\right) + \boldsymbol{\epsilon}\left(\frac{t}{T}\right) \tag{49}$$

或由一个 tvSTARMA($1_1, 1_1$) 过程生成

$$\begin{aligned}
\mathbf{z}\left(\frac{t}{T}\right) &= \phi_{10}(t)W^{(0)}\mathbf{z}\left(\frac{t-1}{T}\right) + \phi_{11}(t)W^{(1)}\mathbf{z}\left(\frac{t-1}{T}\right)\\
&+ \theta_{10}(t)W^{(0)}\boldsymbol{\epsilon}\left(\frac{t-1}{T}\right) + \theta_{11}(t)W^{(1)}\boldsymbol{\epsilon}\left(\frac{t-1}{T}\right) + \boldsymbol{\epsilon}\left(\frac{t}{T}\right), && (50)
\end{aligned}$$

其中， $\boldsymbol{\epsilon}(t)$ 是均值为零、协方差矩阵为$\mathbf{I}_{nT}\sigma^2$的多元正态过程，且 $\sigma^2 = 1$。空间权重矩阵 W 的元素定义如下：

$$w_{ij} = \frac{d_{ij}^{-0.5}}{\sum_{k\neq i} d_{ik}^{-0.5}}, \tag{51}$$

d_{ij} 是两个位置 $\mathbf{x}_i$ 和 $\mathbf{x}_j$之间的大圆距离（正球面距离或球面距离）。它是球面上两点之间的最短距离。设 $\mathbf{x}_i = (\xi_i, \kappa_i)$ 和 $\mathbf{x}_j = (\xi_j, \kappa_j)$ 为地理位置，其中 ξ_η 和 $\kappa_\eta, \eta = i, j$, 分别为纬度和经度。距离 d_{ij} 由下式给出

$$d_{ij} = R\sigma, \tag{52}$$

其中 R 是地球的半径，并且

$$\sigma = \arccos[\sin\xi_i \sin\xi_j + \cos\xi_i \cos\xi_j \cos(|\kappa_i - \kappa_j|)]. \tag{53}$$

[3.] 使用第 3.3 节和第 4.2 节所述的估计程序，计算 $\phi_{10}(t)$ $\phi_{11}(t)$ 和/或 $\theta_{10}(t)$ $\theta_{11}(t)$ 的估计值。

[4.] 计算预测变量$\hat{\mathbf{z}}$ 的均方误差（均方误差），

$$\text{MSE} = \frac{1}{nT}\sum_{i=1}^{n}\sum_{t=1}^{T}[z_i\left(\frac{t}{T}\right) - \hat{z}_i\left(\frac{t}{T}\right)]^2, \tag{54}$$

其中，$\hat{z}_i\left(\frac{t}{T}\right)$ 是 $z_i\left(\frac{t}{T}\right)$的估计量，即位置 i 和时间 t处的观测值。

5.2 时变参数的表述

将使用两组参数来生成样本数据，即

(i) 第1组:

$$\phi_{10}(t) = 0.5 - \frac{\sin\left(\frac{2\pi t}{T}\right)}{4}, \tag{55}$$

$$\phi_{11}(t) = -0.5 - \frac{\cos\left(\frac{2\pi t}{T}\right)}{4}; \tag{56}$$

(ii) 第2组:

$$\phi_{10}(t) = 0.5\left(1 - \frac{t}{T}\right)^2, \tag{57}$$

$$\phi_{11}(t) = -0.5\left(1 - \frac{t}{T}\right)^2, \tag{58}$$

$$\theta_{10}(t) = 0.5\left(\frac{t}{T}\right)^2, \tag{59}$$

$$\theta_{11}(t) = -0.5\left(\frac{t}{T}\right)^2. \tag{60}$$

所模拟的采样 tvSTAR(1_1)和时变时空自回归移动平均模型($1_1, 1_1$) 过程中的一个位置序列呈现在图 2 中。

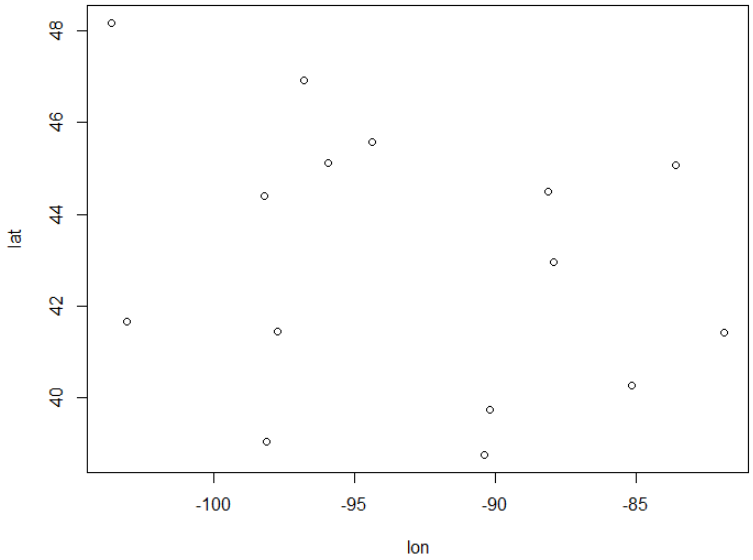
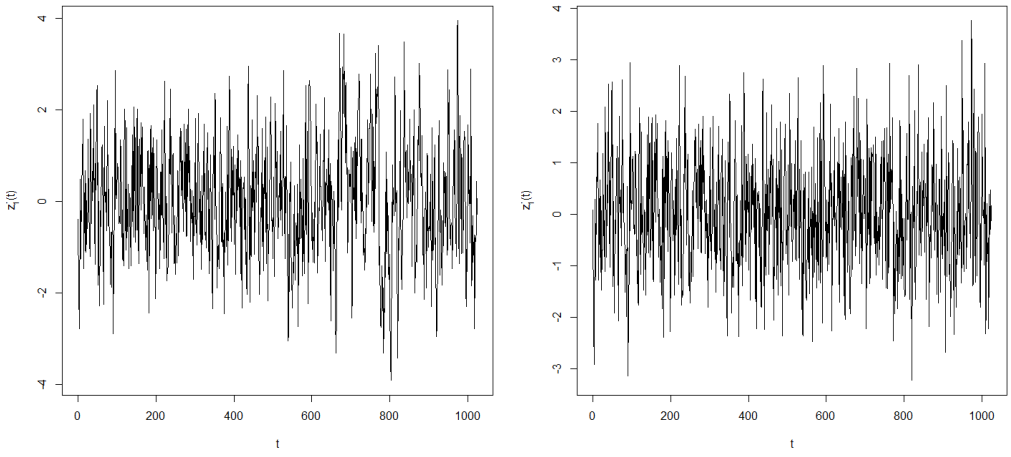


图 1：模拟位置。



(a) 第1组　　(b) 第2组

图 2：一个位置的模拟序列。

5.3 结果

由于真实参数是平滑函数，因此选择了墨西哥帽小波。请注意，所有估计值均使用 $J = 2$ 计算。

图3显示了**z**估计值的均方误差箱线图。可以看出，两组均方误差的前三个四分位数大致相同，两组都有离群值，且第2组的变异性较低。

图4和图5分别展示了两组参数真实值与1000次实验中估计值均值的比较

分别对应第1组和第2组。我们看到，两种情况下拟合效果都很好。值得注意的是，第1组的估计值可能更好，这是因为我们采用了时变 STAR 模型；而第2组则纳入了 MA 项。总之，对于这两组，所提出的估计程序都表现出色。

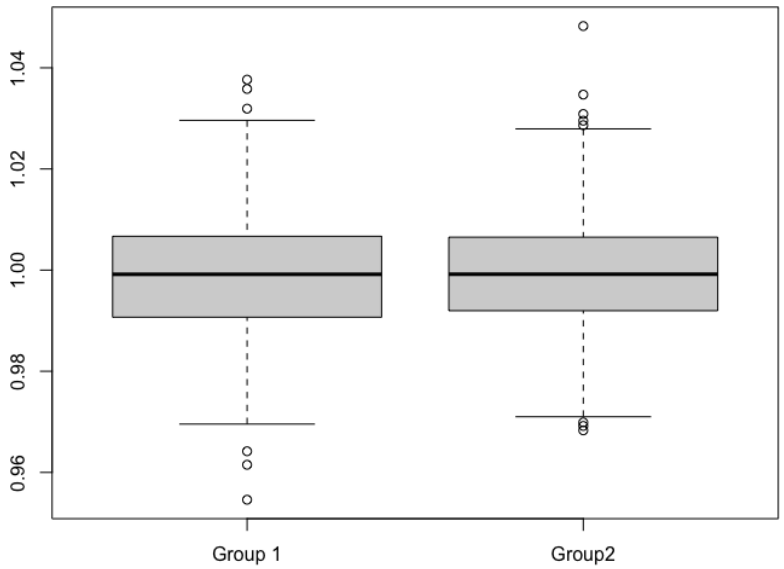


图 3：1000 次实验的均方误差箱线图。

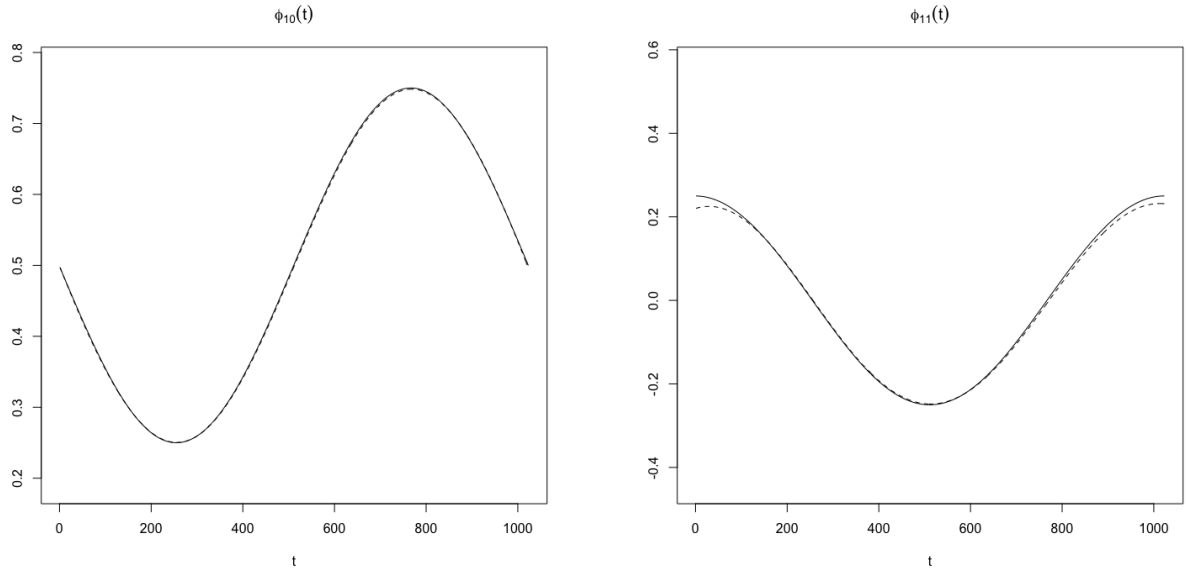


图4：第1组的真实参数（实线）与通过墨西哥帽小波获得的估计值平均值（点线）的比较。

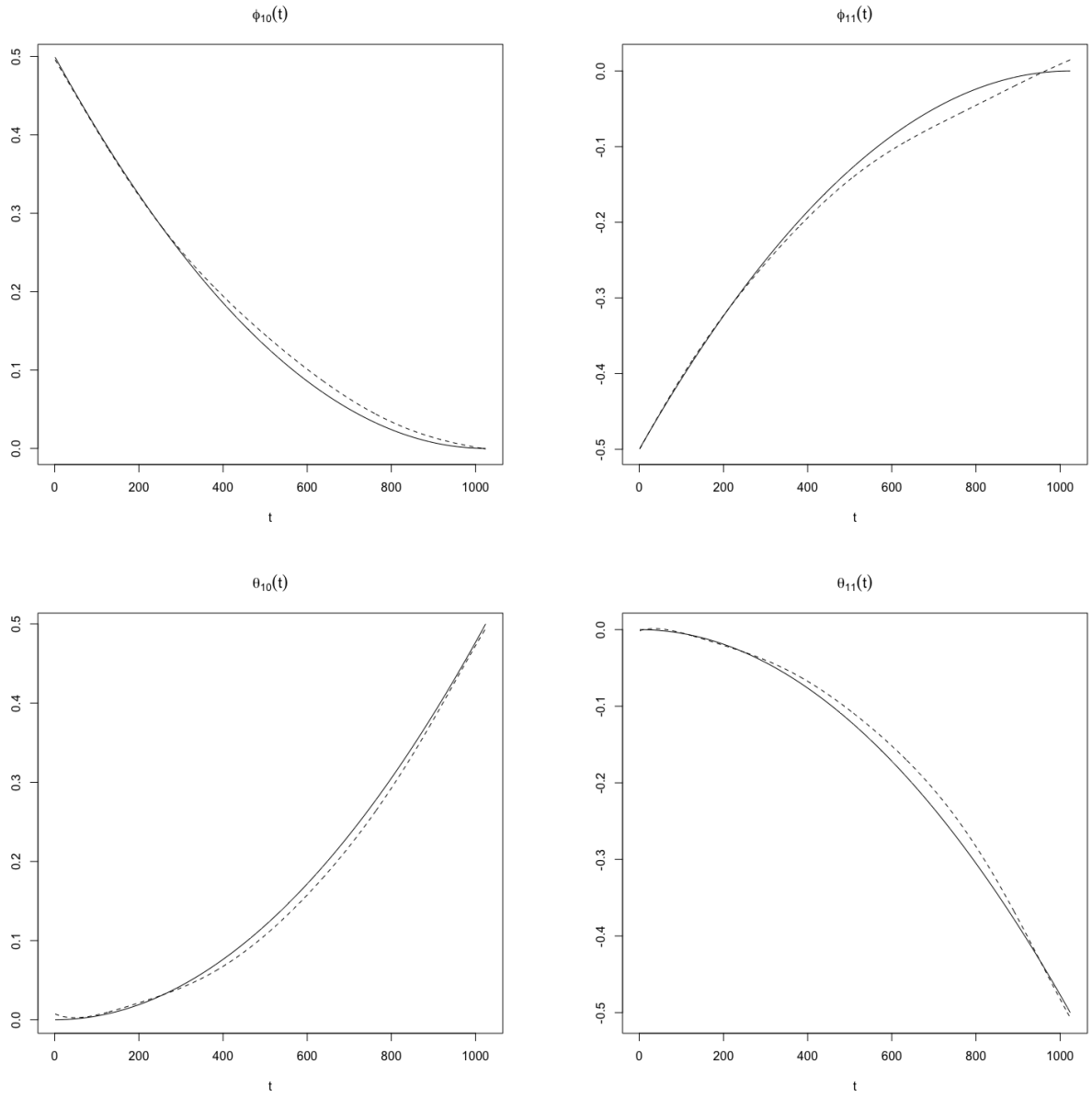


图 5：第2组真实参数（实线）与通过墨西哥帽小波获得的估计值均值（虚线）的比较。

6 通过模拟比较空间权重矩阵

受 Jin（2017）的启发，本节给出一些模拟示例，以比较不同类型的空间权重矩阵。如第 3.1 节所述，在 tvSTARMA模型中，空间权重矩阵也是一个主要特征。矩阵 W 的距离有多种选择，例如逆距离、负指数距离、 k 近邻等。我们的模拟选用了逆权重距离和负指数权重距离。它们定义如下

$$w_{ij}^{di} = \begin{cases} \frac{d_{ij}^{-\alpha}}{\sum_{k \neq i} d_{ik}^{-\alpha}}, & i \neq j, \\ 0, & i = j \end{cases} \tag{61}$$

and

$$w_{ij}^{ne} = \begin{cases} \frac{\exp\{-\alpha d_{ij}\}}{\sum_{k\neq i}\exp\{-\alpha d_{ik}\}}, & i \neq j, \\ 0, & i = j, \end{cases} \tag{62}$$

其中， d_{ij} 是大圆距离，二者都取决于参数 α，该参数为正数。

6.1 模拟步骤

模拟包括以下步骤：

[1.]我们考虑如图 1 所示生成的 $n = 15$ 个样本位置，然后模拟 n 个长度为 $T = 1024$的时间序列实验。

[2.] 样本数据 $\mathbf{z} = [\mathbf{z}(\frac{1}{T}) ..., \mathbf{z}(\frac{T}{T})]$ 由一个均值函数为 $\mu(\mathbf{z}) = \mathbf{0}$、协方差函数为以下形式的高斯分布随机场模拟得到

$$\mathrm{Cov}(z_{\mathbf{x}+\mathbf{h}}(t+\tau), z_{\mathbf{x}}(t)) = \frac{\sigma^2}{\beta(|\tau|)^{d/2}}\omega\left\{\frac{||\mathbf{h}||}{\sqrt{\beta(|\tau|)}}\right\} + \sigma_\epsilon^2 \mathbb{I}_{(z_{\mathbf{x}+\mathbf{h}}(t+\tau)=z_{\mathbf{x}}(t))}, (\mathbf{h};\tau) \in R^d \times R, \tag{63}$$

其中，$\mathbf{h}$ 和 τ 分别为空间和时间偏移， σ^2 =Var($z_{\mathbf{x}}(t)$)， $d = 2$ 和 $\sigma_\epsilon^2 = 0.05$； $\omega(\cdot) \geq 0$ 是完全单调函数，而 $\beta(\cdot)$ 是导数完全单调的正函数，定义为

$$\beta(f) = (f^\zeta + 1)^{\delta/\zeta}, \tag{64}$$

$$\omega(f) = \exp\{-f/\gamma\}, \tag{65}$$

其中， $\zeta = 1$、 $\delta = 0.5, 1, 1.5$ 和 $\gamma = 0.25, 0.5, 1$ 已选定。

[3.] 使用第 3.3 节和第 4.2 节所述的估计程序，采用不同的空间权重矩阵拟合 tvSTAR(1_1) 和时变时空自回归移动平均模型($1_1, 1_1$)，其中反距离权重距离取 $\alpha = 0.5$、1，负指数权重距离取 $\alpha = 0.5$、1、2。

[4.] 计算预测变量 $\hat{\mathbf{z}}$ 的均方误差 (MSE)，

$$\mathrm{MSE} = \frac{1}{MnT}\sum_{m=1}^{M}\sum_{i=1}^{n}\sum_{t=1}^{T}[z_i^m\left(\frac{t}{T}\right) - \hat{z}_i^m\left(\frac{t}{T}\right)]^2, \tag{66}$$

其中$\hat{z}_i^m\left(\frac{t}{T}\right)$ 是 $z_i^m\left(\frac{t}{T}\right)$的估计量，即第 m 次实验中位置 i、时间 t 的观测值

t.

6.2 结果

根据表 1–3，可以得出：(i) 对于任意参数取值，J 较大时，均方误差略有下降；(ii) 对于任意参数取值，随着 α 增大，均方误差会下降；(iii) 当 $\delta = 0.5$ 且 γ 取任意值时，均方误差较小；(iv) 时变时空自回归移动平均模型($1_1, 1_1$)的均方误差较小；(v) 在几乎所有情况下，采用反距离和 $\alpha = 1$ 的矩阵均能得到较小的均方误差。在这些情况下，空间权重矩阵的影响相似。

7 应用

降水是气候、水文气象学研究及农业生产中的重要变量。许多研究使用时间序列模型分析不同区域的降水，例如 Dalezios 和 Adamowski（1995）、Wang 等（2013）以及 Wu 等（2021）。本文使用通过 R 软件包 rnoaa 从以下来源获取的历史日降水记录：

表 1: 采用墨西哥帽小波和 $J = 2$ 拟合的 tvSTAR(1_1) 模型在不同数据集上的均方误差 .

空间权重矩阵	w_{ij}^{di}		w_{ij}^{ne}		
	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$
$\gamma = 0.25$ $\delta = 0.5$	0.5720284	0.5720198	0.5720292	0.5720263	**0.5720187**
$\gamma = 0.25$ $\delta = 1$	0.8099595	**0.8099531**	0.8099612	0.8099605	0.8099591
$\gamma = 0.25$ $\delta = 1.5$	0.9289999	**0.9289853**	0.9290054	0.9290036	0.9290000
$\gamma = 0.5$ $\delta = 0.5$	0.5719711	**0.5718976**	0.5719838	0.5719739	0.5719395
$\gamma = 0.5$ $\delta = 1$	0.8097927	**0.8096183**	0.8098543	0.8098237	0.8097445
$\gamma = 0.5$ $\delta = 1.5$	0.9287466	**0.9284917**	0.9288502	0.9287975	0.9286744
$\gamma = 1$ $\delta = 0.5$	0.5713549	**0.5709173**	0.5715235	0.5714033	0.5711274
$\gamma = 1$ $\delta = 1$	0.8084149	**0.8075349**	0.8088334	0.8085280	0.8079301
$\gamma = 1$ $\delta = 1.5$	0.9267936	**0.9256779**	0.9272953	0.9268617	0.9260850

表 2: 均 采用墨西哥帽小波和 $J =$ 拟合的 tvSTAR(1_1) 模型在不同数据集上的均方误差 3.

空间权重矩阵	w_{ij}^{di}		w_{ij}^{ne}		
	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$
$\gamma = 0.25$ $\delta = 0.5$	0.5713301	0.5713054	0.5713330	0.5713241	**0.5713024**
$\gamma = 0.25$ $\delta = 1$	0.8092726	**0.8092524**	0.8092760	0.8092697	0.8092562
$\gamma = 0.25$ $\delta = 1.5$	0.9283331	**0.9283079**	0.9283398	0.9283333	0.9283203
$\gamma = 0.5$ $\delta = 0.5$	0.5712721	**0.5711897**	0.5712766	0.5712649	0.5712213
$\gamma = 0.5$ $\delta = 1$	0.8090920	**0.8089110**	0.8091462	0.8091135	0.8090265
$\gamma = 0.5$ $\delta = 1.5$	0.9280607	**0.9278012**	0.9281587	0.9281040	0.9279749
$\gamma = 1$ $\delta = 0.5$	0.5706313	**0.5702010**	0.5707714	0.5706620	0.5703890
$\gamma = 1$ $\delta = 1$	0.8076571	**0.8067883**	0.8080496	0.8077554	0.8071624
$\gamma = 1$ $\delta = 1.5$	0.9259958	**0.9249419**	0.9265289	0.9261045	0.9253332

GHCN（全球历史气候学网络）每日数据集，这是美国国家海洋和大气管理局（NOAA）的综合公共数据库。所选数据为美国中西部各州的日降水量（单位为十分之一毫米）。该区域由 11 个州组成：北达科他州、南达科他州、伊利诺伊州、艾奥瓦州、堪萨斯州、密歇根州、明尼苏达州、密苏里州、内布拉斯加州、俄亥俄州和威斯康星州。我们仅使用在 1990-01-01 至 2009-12-30（含首尾日期）期间没有缺失数据的气候监测站点。所选的 30 个站点见图6，图7展示了两个气象站记录的降水量。

由于降水数据可能因存在概率非零的极大值而呈现重尾分布，我们应用变换

$$Z_i\left(\frac{t}{T}\right) = \log_{10}\left(Y_i\left(\frac{t}{T}\right) + 1\right), \tag{67}$$

其中， $Y_i\left(\frac{t}{T}\right)$ 是原始的 位置 $\mathbf{x}_i$ 和时间 t 序列，然后应用所提出的估计方法来消除这些影响。

表3：均方误差 o　采用墨西哥帽小波的拟合时变时空自回归移动平均模型$(1_1, 1_1)$模型的 $J = 2$个不同数据集 f　　.

空间权重矩阵	w_{ij}^{di}		w_{ij}^{ne}		
	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$
$\gamma = 0.25$ $\delta = 0.5$	0.5493941	**0.5493536**	0.5494022	0.5493865	0.5493567
$\gamma = 0.25$ $\delta = 1$	0.7960634	**0.7960057**	0.7960567	0.7960479	0.7960121
$\gamma = 0.25$ $\delta = 1.5$	0.9221163	**0.9220534**	0.9221938	0.9223765	0.9221010
$\gamma = 0.5$ $\delta = 0.5$	0.5493746	**0.5492904**	0.5493905	0.5493671	0.5493187
$\gamma = 0.5$ $\delta = 1$	0.7960146	**0.7957764**	0.7963313	0.7960042	0.7958850
$\gamma = 0.5$ $\delta = 1.5$	0.9217563	**0.9214193**	0.9218807	0.9217912	0.9215954
$\gamma = 1$ $\delta = 0.5$	0.5490224	**0.5487275**	0.5491266	0.5490394	0.5488448
$\gamma = 1$ $\delta = 1$	0.7948874	**0.7941313**	0.7952156	0.7949564	0.7944082
$\gamma = 1$ $\delta = 1.5$	0.9200001	**0.9189014**	0.9204669	0.9200260	0.9192314

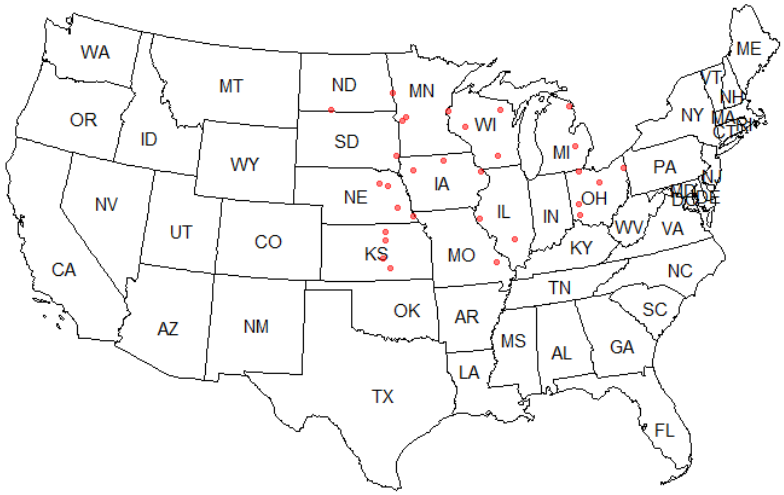


图6：从美国中西部各州选出的站点的位置。

本应用选用逆权重距离和负指数权重距离。权重定义为（61）和（62）。对于逆权重和负指数权重，我们分别设置 $\alpha = 0.5$,1 和 0.5、1、2。

这些结果是使用第 3.2 节和第4.2节所述的估计程序，通过 tvSTAR(1_1) 和时变时空自回归移动平均模型$(1_1, 1_1)$模型获得的。时变参数的小波展开使用墨西哥帽小波以及 $J = 2$、3、4 构建（最后一个值仅用于 tvSTAR 模型）。

表 4 显示了均方误差。 不同空间权重矩阵下各模型估计值的均方误差。我们

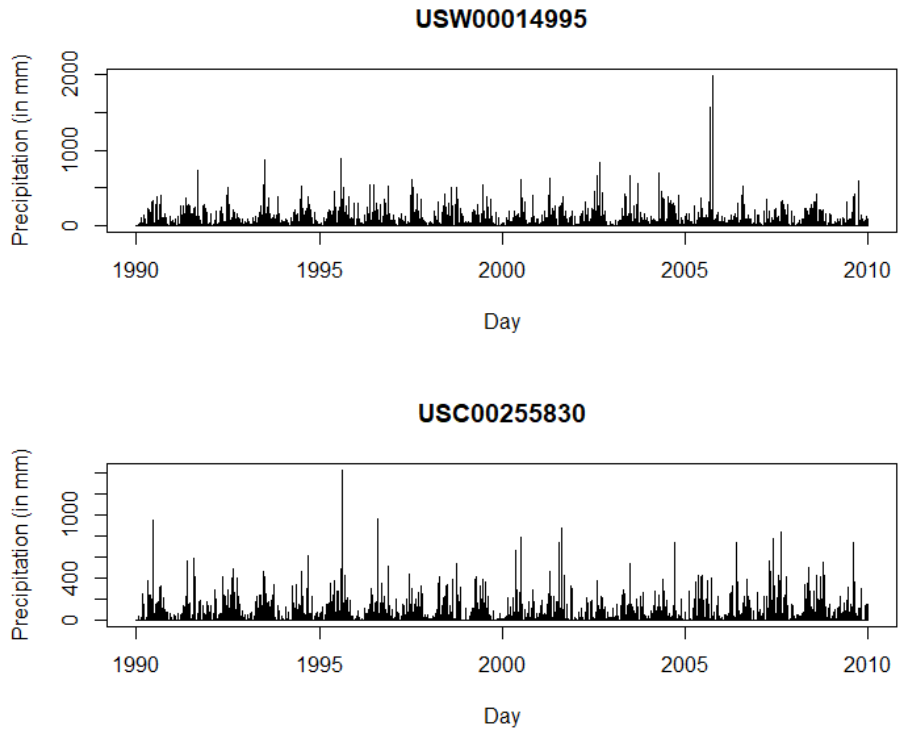


图 7：2 个站点记录的降水。

可以看到，当空间权重矩阵采用逆距离权重并结合 $\alpha = 1$ 时，均方误差更小。请注意，在相同的拟合模型下，均方误差相近（与模拟中的结论相同），因此 $J = 2$ 可能是拟合数据的合适选择，因为它的系数较少。

表 4：使用墨西哥帽小波的不同模型的均方误差。

空间权重矩阵	w_{ij}^{di}		w_{ij}^{ne}		
	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$
tvSTAR(1_1) $J = 2$	1.142290e-4	**1.140377e-4**	1.148335e-4	1.143726e-4	1.141951e-4
tvSTAR(1_1) $J = 3$	1.142095e-4	**1.140122e-4**	1.148155e-4	1.143552e-4	1.141754e-4
tvSTAR(1_1) $J = 4$	1.141900e-4	**1.139864e-4**	1.147969e-4	1.143357e-4	1.141545e-4
时变时空自回归移动平均模型($1_1, 1_1$) $J = 2$	1.141545e-4	**1.139164e-4**	1.147790e-4	1.142976e-4	1.140730e-4
时变时空自回归移动平均模型($1_1, 1_1$) $J = 3$	1.141271e-4	**1.138722e-4**	1.147541e-4	1.142734e-4	1.140426e-4

8 进一步的评论

平稳性是时间序列分析中的一项基本假设，但在实践中有时难以保证。本文基于 Dahlhaus 意义下的局部平稳过程这一概念，提出了时变 STARMA 模型。时变参数以小波基展开。我们进行了若干模拟，并

采用所提出的估计程序给出了一个应用实例。我们选择了墨西哥帽小波，但也可以使用其他基，例如莫莱小波、香农小波或迈耶小波。对于时变 STARMA 模型，未来研究中有许多值得探讨的问题，例如：(i) 纳入季节性因素；(ii) 探究创新协方差函数异质性；(iii) 考虑空间权重矩阵随时间变化的情形；(iv) 将所提出的模型与其他时空模型进行比较，例如空间动态模型和层次时空模型；(v) 研究 tvSTARMA 模型的渐近性质，如命题 1 和 2 对 tvSTAR 模型所做的研究。

致谢

Yangyang Chen 感谢圣保罗州研究基金会通过资助项目 2019/05917-6 提供支持。Pedro Alberto Morettin 和 Chang Chiann 感谢圣保罗州研究基金会资助项目 2018/04654-9 提供的部分支持。

A 附录

A.1 命题1的证明

公式 (23) 可写为

$$\boldsymbol{\nu}\left(\frac{t}{T}\right)=\sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}R_{sl}\left(\frac{t}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{t-s}{T}\right)+\boldsymbol{\epsilon}\left(\frac{t}{T}\right), \tag{A.1}$$

其中

$$R_{sl}\left(\frac{t}{T}\right)=\sum_{j\geq J}^{\infty}\sum_{m=0}^{2^j-1}\beta_{j,m}^{sl}\psi_{j,m}\left(\frac{t}{T}\right). \tag{A.2}$$

定义

$$S=\left[\sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}R_{sl}\left(\frac{p+1}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{p+1-s}{T}\right),...,\sum_{s=1}^{p}\sum_{l=0}^{\lambda_s}R_{sl}\left(\frac{T-p}{T}\right)W^{(l)}\mathbf{Z}\left(\frac{T-p-s}{T}\right)\right]' \tag{A.3}$$

and

$$\boldsymbol{\epsilon}=\left[\mathrm{vec}\left(\boldsymbol{\epsilon}\left(\frac{p+1}{T}\right)\right)',...,\mathrm{vec}\left(\boldsymbol{\epsilon}\left(\frac{T-p}{T}\right)\right)'\right]', \tag{A.4}$$

方程（29）可分解为

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= (\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}'(\boldsymbol{\Psi}\boldsymbol{\beta}+S+\boldsymbol{\epsilon})\\ &= \boldsymbol{\beta}+(E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}'\boldsymbol{\epsilon}+[(\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}-(E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}]\boldsymbol{\Psi}'\boldsymbol{\epsilon}+(\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}'S\\ &= \boldsymbol{\beta}+T_1+T_2+T_3. \end{aligned} \tag{A.5}$$

由于 $T_1=(E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}'\boldsymbol{\epsilon}$，我们有 $E[T_1]=0$，并对矩阵 $(\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}$ 进行泰勒展开，得到 $T_2=T_{21}+T_{22}$，其中

$$T_{21}=(E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}(E\boldsymbol{\Psi}'\boldsymbol{\Psi}-\boldsymbol{\Psi}'\boldsymbol{\Psi})(E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}'\boldsymbol{\epsilon} \tag{A.6}$$

and

$$T_{22}=(E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}(E\boldsymbol{\Psi}'\boldsymbol{\Psi}-\boldsymbol{\Psi}'\boldsymbol{\Psi})(E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}(E\boldsymbol{\Psi}'\boldsymbol{\Psi}-\boldsymbol{\Psi}'\boldsymbol{\Psi})(E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}'\boldsymbol{\epsilon}. \tag{A.7}$$

类似于 Chang 和 Morettin（2005），我们得到

$$\|T_{21}\|_2 \leq \left\|(E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\right\|_2^2\left\|\mathrm{Cov}(\boldsymbol{\Psi}'\boldsymbol{\Psi})\right\|_2^{1/2}\left\|\mathrm{Cov}(\boldsymbol{\Psi}'\boldsymbol{\epsilon})\right\|_2^{1/2}$$

$$
\begin{aligned}
&= O((nT)^{-2})O(2^{2J}(nT)^{1/2})O((nT)^{1/2}) \\
&= O(2^{J}(nT)^{-1}) \qquad \text{(A.8)}
\end{aligned}
$$

and

$$
\begin{aligned}
\|T_{22}\|_2 &\leq \left\|(E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\right\|_2^3 \left\|\mathrm{Cov}(\boldsymbol{\Psi}'\boldsymbol{\Psi})\right\|_2^2 \left\|\mathrm{Cov}(\boldsymbol{\Psi}'\boldsymbol{\epsilon})\right\|_2 \\
&= O((nT)^{-3})O(2^{2J}nT)O((nT)^{1/2}) \\
&= O(2^{2J}(nT)^{-3/2}). \qquad \text{(A.9)}
\end{aligned}
$$

根据 Donoho 等人（1995），我们有

$$
\sup\left\{\sum_{j\geq J}^{\infty}\sum_{k=0}^{2^j-1}|\beta_{j,k}^{sl}|^2\right\} = O(2^{-2Ju}), \qquad \text{(A.10)}
$$

其中 $u = d + 1/2 - 1/\tilde{p}$ 满足 $\tilde{p} = \min\{p, 2\}$，并且与 Dahlhaus 等人（1999）的做法类似，

$$
\begin{aligned}
\|T_3\|_2 &\leq (\left\|(E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\right\|_2 + \left\|(\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1} - (E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\right\|_2)\left\|\boldsymbol{\Psi}'S\right\|_2 \\
&= (O((nT)^{-1}) + O(2^{J}(nT)^{-3/2}) + O(2^{2J}(nT)^{-2}))\left\|\boldsymbol{\Psi}'S\right\|_2 \\
&= O((nT)^{-1})O(nT(2^{-Ju} + (nT)^{-1/2}2^{-J(d-1/2-1/(2\tilde{p}))})\sqrt{\log(nT)})) \\
&= O((2^{-Ju} + (nT)^{-1/2}2^{-J(d-1/2-1/(2\tilde{p}))})\sqrt{\log(nT)})) \\
&= O((nT)^{-1/2-\tau}), \qquad \text{(A.11)}
\end{aligned}
$$

对于某个 $\tau > 0$。于是，$\|T_3\|_2 \leq O((nT)^{-1/2})$。

结论（$i.$）得证。

对于（$ii.$），我们有

$$
\begin{aligned}
E[T_1T_1'] &= (E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\mathrm{Cov}(\boldsymbol{\Psi}'\boldsymbol{\epsilon})(E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1} \\
&\leq \left\|(E\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\right\|_2^2 \left\|\mathrm{Cov}(\boldsymbol{\Psi}'\boldsymbol{\epsilon})\right\|_2 \\
&= O((nT)^2)O(nT) \\
&= O((nT)^{-1}). \qquad \text{(A.12)}
\end{aligned}
$$

同理，

$$
\begin{aligned}
E[T_2T_2'] &= O(2^{2J}(nT)^{-2}) + O(2^{4J}(nT)^{-3}), && \text{(A.13)} \\
E[T_3T_3'] &= O((nT)^{-1}), && \text{(A.14)} \\
E[T_1T_2'] &= E[T_1'T_2] = E[T_2T_3'] = E[T_2'T_3] = O(2^{J}(nT)^{-3/2}) + O(2^{2J}(nT)^{-2}), && \text{(A.15)} \\
E[T_1T_3'] &= E[T_1'T_3] = O((nT)^{-1}), && \text{(A.16)}
\end{aligned}
$$

于是结论（$ii.$）得证。

A.2 命题 2 的证明

引理 1 假设 A.1–A.4 成立，则

$$
N\mathbf{H}(\boldsymbol{\Psi}'\boldsymbol{\Psi})^{-1}\mathbf{H}' \xrightarrow{P} \boldsymbol{\Gamma}, \qquad \text{(A.17)}
$$

其中，$\boldsymbol{\Psi}$ 与式（27）中相同，并且

$$
\boldsymbol{\Gamma} = \lim_{N\to\infty} N\mathbf{H}E[\boldsymbol{\Psi}'\boldsymbol{\Psi}]^{-1}\mathbf{H}'. \qquad \text{(A.18)}
$$

证明。令 $\mathbf{\Psi}$中第 ψ_{gl} 行、第 g列的单元格为 l，则有

$$\psi_{gl} = \sum_{i=1}^{n} f_{gl}^{i} Z_{gl}^{i}, \tag{A.19}$$

其中，f_{gl}^{i} 由一个小波函数与一个权重函数的乘积得到，Z_{gl}^{i} 是一个局部平稳过程，二者都是位于 i处的 $\mathbf{\Psi}$ 的第(g,l)个单元的元素。

设任意 $\mathbf{\Psi}'\mathbf{\Psi}$ 的子分区中第 l 行、第 m 列的单元格 $(\psi\psi)_{lm}$，$(\psi\psi)_{lm}$ 可表示为

$$\begin{aligned}
(\psi\psi)_{lm} &= \sum_{g=1}^{N} \psi_{gl}\psi_{gm} \\
&= \sum_{g=1}^{N} \left[\sum_{i=1}^{n} (f_{gl}^{i} Z_{gl}^{i}) \sum_{i=1}^{n} (f_{gm}^{i} Z_{gm}^{i}) \right] \\
&= \sum_{g=1}^{N} f_{gl}^{1} Z_{gl}^{1} f_{gm}^{1} Z_{gm}^{1} + \sum_{g=1}^{N} f_{gl}^{1} Z_{gl}^{1} f_{gm}^{2} Z_{gm}^{2} + \cdots + \sum_{g=1}^{N} f_{gl}^{n} Z_{gl}^{n} f_{gm}^{n} Z_{gm}^{n}.
\end{aligned}$$

$(\psi'\psi)_{lm}$ 的证明与 Sato 等人（2007年）的证明类似，且该结果由 Andrews（1988年）的 L^1 $-Maxingales$ 弱大数定律得出。

引理 2 假设 A.1-A.4 成立，则

$$\frac{1}{\sqrt{N}} \mathbf{G}(\mathbf{\Psi}'\boldsymbol{\epsilon}) \xrightarrow{D} \mathcal{N}_K(\mathbf{0}, \mathbf{\Gamma}^{-1}), \tag{A.20}$$

其中 $\boldsymbol{\epsilon}$ 如式 (A.4) 所示，且为协方差矩阵。

$$\mathbf{\Gamma}^{-1} = \lim_{N\to\infty} \mathbf{G} \frac{E[\mathbf{\Psi}'\mathbf{\Psi}]}{N} \mathbf{G}', \tag{A.21}$$

其中 $N = n(T-p)$，$\mathbf{G}$ 是一个与 $\mathbf{\Psi}$ 具有相同列数和 K 行数的矩阵 .

证明。令 $(\psi e)_l$ 为 $\mathbf{\Psi}'\boldsymbol{\epsilon}$ 的第 l 个元素，

$$\begin{aligned}
(\psi e)_l &= \sum_{g=1}^{N} (\psi e)_{gl} \\
&= \sum_{g=1}^{N} f_{gl} \tilde{Z}_{gl} \epsilon_g,
\end{aligned}$$

˜其中，f_{gl} 是一个小波函数，Z_{gl} 是由 n 个局部平稳过程按权重加权的加权算术平均，且权重之和为一；二者均属于 $\mathbf{\Psi}$ 的第 $(g,\ l)$ 个单元，并且 ϵ_g 的第 g 个元素 $\boldsymbol{\epsilon}$ 是独立高斯过程。

令 $\mathbf{\Psi}'\boldsymbol{\epsilon} = \sum_{g=1}^{N} b_g$，其中 b_g 是由 $(\psi e)_{gl}$ 构成的向量，并将 $\mathscr{F}_g$ 定义为包含截至时刻 g 所有信息的 σ-代数。由于

$$E[b_g \mid \mathscr{F}_{g-1}] = 0, \tag{A.22}$$

$\{b_g\}_{g=1}^{\infty}$ 是一个鞅差序列。于是

$$\begin{aligned}
E\left[\frac{\sum_{g=1}^{N} b_g b_g'}{N}\right] &= E\left[\frac{\mathbf{\Psi}'\boldsymbol{\epsilon}\boldsymbol{\epsilon}'\mathbf{\Psi}}{N}\right] \\
&= E\left[E\left[\frac{\mathbf{\Psi}'\boldsymbol{\epsilon}\boldsymbol{\epsilon}'\mathbf{\Psi}}{N} \,\middle|\, \mathbf{\Psi}\right]\right] \\
&= E\left[\frac{\mathbf{\Psi}'\mathbf{\Psi}}{N}\right].
\end{aligned}$$

类似地，

$$\mathbf{G}E\left[\frac{\sum_{g=1}^{N} b_g b_g'}{N}\mathbf{G}' = \mathbf{G}E\left[\frac{\mathbf{\Psi}'\mathbf{\Psi}}{N}\right]\right]\mathbf{G}' \to \mathbf{\Gamma}^{-1}, \tag{A.23}$$

其中

$$\mathbf{\Gamma}^{-1} = \lim_{N\to\infty} \mathbf{G}\frac{E[\mathbf{\Psi}'\mathbf{\Psi}]}{N}\mathbf{G}' \tag{A.24}$$

并且，与引理 1 类似，

$$\mathbf{G}\frac{\sum_{g=1}^{N} b_g b_g'}{N}\mathbf{G}' \xrightarrow{P} \mathbf{\Gamma}^{-1}. \tag{A.25}$$

该结果由 White（2000）关于鞅差序列的中心极限定理得出。

由公式 (A.11)，$\sqrt{N}\|T_3\|_2 = O((nT)^{-\tau}) = o_p(1)$，应用引理 1、引理 2 和 Slutsky 定理，可得

$$\begin{aligned}
\sqrt{N}\mathbf{H}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) &= \sqrt{N}\mathbf{H}(T_1 + T_2 + T_3) && \text{(A.26)}\\
&= \sqrt{N}\mathbf{H}(T_1 + T_2) + o_p(1) && \text{(A.27)}\\
&= \sqrt{N}\mathbf{H}(\mathbf{\Psi}'\mathbf{\Psi})^{-1}\mathbf{\Psi}'\boldsymbol{\epsilon} + o_p(1) && \text{(A.28)}\\
&= \mathscr{N} + o_p(1), && \text{(A.29)}
\end{aligned}$$

其中，$\mathscr{N}$服从均值为零、协方差矩阵为 (A.18) 中所示 $\mathbf{\Gamma}$ 的 K 维正态分布。该结果由 Slutsky 定理得出。

参考文献

1. Andrews, D. W. K. (1988年) 相关且非同分布随机变量的大数定律，*Econometric Theory*,**4**(3), 458-467. 2. Box, G. E. P. and Jenkins, G. M. (1970) *Time series Analysis: Forecasting and Control*, 旧金山：Holden-Day. 3. Chiann, C. and Morettin, P. A. (1999) 时变线性系统的估计，*Statistical Inference for StochasticProcesses* **2**, 253–285. 4. Chiann, C. and Morettin, P. A. (2005) 时变线性系统的时域非线性估计，*Journalof Nonparametric Statistics*,**17**(3), 365–383. 5. Cipra, T. and Motykova, I. (1987) 时间序列分析中的卡尔曼滤波器研究，*Commentationes Mathematicae Universitatis Carolinae*,**28**(3):549-563. 6. Cliff, A. D. and Ord, J. K. (1975) 时空建模及其在区域预测中的应用，*Transactionsof the Institute of British Geographers*, 第64期, 119-128. 7. Dahlhaus, R. (1996a) 局部平稳过程的最大似然估计与模型选择，*Journalof Nonparametric Statistics*, **6**(2-3), 171-191. 8. Dahlhaus, R. (1996b) 局部平稳过程的 Kullback–Leibler 信息散度，*StochasticProcesses and their Applications*,**62**(1), 139–168. 9. Dahlhaus, R. (1996c) 具有演化谱的非平稳过程的渐近统计推断，载于：Robinson, P. M., Rosenblatt, M. (编) *Athens Conference on Applied Probability and Time Series Analysis. Volume II: Time Series Analysis In Memory of E.J. Hannan*, **115**, 145–159. Springer, 纽约州纽约市。10. Dahlhaus, R. (1997) 将时间序列模型拟合到非平稳过程，*The Annals of Statistics*, **25**(1), 1–37.

11. Dahlhaus, R. (2000) 局部平稳过程的似然近似， *The Annals of Statistics*, 28(6), 1762–1794. 12. Dahlahus, R. (2012) 局部平稳过程， *Time Series Analysis: Methods and Applications*, 30, 351-413. 13. Dahlhaus, R., Neumann, M. H. and von Sachs, R. (1999) 时变自回归过程的非线性小波估计， *Bernoulli* 5(5), 873–906. 14. Dalezios, N. R and Adamowski, K. (1995) 农村流域的时空降水建模， *HydrologicalSciences Journal*, 40(5), 553-568. 15. Donoho, D., Johnstone, I., Kerkyacharian, G., Picard, D. (1995) 小波收缩：渐近极限？ *Journal of theRoyal Statistical Society*, B 辑（方法论）, 57(2), 301-337. 16. Hu, X., Qin, Z. and Chu, F. (2011) 基于时空自回归移动平均过程的板结构损伤检测, *Journal of Physics: Conference Series*, 305(1):012119. 17. Jin, E. Y. (2017) 时空 STARMA 模型中空间邻接结构，硕士论文，圣保罗大学数学与统计研究所，圣保罗。18. Kalman, R. (1960年) 线性滤波与预测问题的新方法, *Journal of Basic Engineering*, 82(1), 35-45. 19. Kamarianakis, Y. and Prastacos, P. (2005) 交通流的时空建模, *Computers & Geosciences*, 31(2), 119-133. 20. Kurt, S. and Tunay, K. B. (2015) 基于卡尔曼滤波器的 STARMA 模型估计：区域银行存款案例。 *Procedia - Social and Behavioral Sciences*, 第195卷, 2537-2547. 21. Martin, R. L. and Oeppen, J. E. (1975) 利用时空相关函数识别区域预测模型， *Transactions of the Institute of British Geographers*, 第66期, 95-118. 22. Pace, R. K., Barry, R., Gilley, O. W. and Sirmans, C. F. (2000) 一种时空预测方法及其在房地产价格中的应用， *International Journal of Forecasting*, 16(2), 229-246. 23. Pfeifer, P. E., and Deutsch, S. J. (1980a) 时空建模的三阶段迭代程序， *Technometrics*, 22(1), 35-47. 24. Pfeifer, P. E., and Deutsch, S. J. (1980b) 一阶时空 ARMA 模型的识别与解释， *Technometrics*, 22(3), 397-408. 25. Pfeifer, P. E., and Deutsch, S. J. (1981a) 同时相关变量的样本时刻自相关函数的方差， *SIAM Journal of Applied Mathematics*, 40(1), 133-136. 26. Pfeifer, P. E., and Deutsch, S. J. (1981b) 季节性时空 ARIMA 建模， *Geographical Analysis*, 13(2), 117-133. 27. Pfeifer, P. E., and Deutsch, S. J. (1981c) 同时相关创新项的时空 ARMA 建模， *Technometrics*, 23(4), 410-409. 28. Rohan, N. and Ramanathan T. V. (2013) 时变 GARCH 模型的非参数估计， *Journal ofNonparametric Statistics*, 25(1), 33-52.

29. Sato, J. R., Morettin, P. A., Arantes, P. R. 和 Amaro Jr, E.（2007）基于小波的时变向量自回归建模，*Computational Statistics & Data Analysis*,**51**(12), 5847-5866。30. Wang, S., Feng, J. 和 Liu, G.（2013）季节性时间序列模型在降水预报中的应用，*Mathematical and Computer Modelling*, **58**(3–4), 677-683。31. White, H.，（2000）《计量经济学家的渐近理论》，修订版，Academic Press，纽约。32. Wu, X., Zhou, J., Yu, H., Liu, D., Xie, K., Chen, Y., Hu, J., Sun, H. 和 Xing, F.（2021）用于降水量预测与干旱分析的混合小波-ARIMA-LSTM 模型的开发， *Atmosphere*, **12**(1):74。33. Yousuf, K. 和 Ng, S.（2021）利用时变参数增强高维预测回归，*Journal of Econometrics*,**224**(1), 60-87。34. Zou, J., Zhu, J., Xie, P., Xuan, P. 和 Lai, X.（2018）风力发电时空序列的时空自回归移动平均模型。*2018 IEEE Power & Energy Society General Meeting (PESGM)*, 1-5。