Supplementary Material for: "The Arrow of Time in Multivariate Time Series"

A. Vectorized Darmois-Skitovich theorem

The Darmois-Skitovich theorem is provided below as Lemma A.1. It is used for proving the identifiability of independent component analysis (Comon, 1994) and lies at the heart of the proof of Theorem 2.2.

Lemma A.1 (Vectorised Darmois-Skitovich theorem for infinite sums, Theorem 4 in (Ibragimov, 2014))**.** *Let* $\mathbf{X}_1, \mathbf{X}_2, \ldots$ *be independent d-dimensional random vectors and consider the linear combinations* $L_1 = \sum_{j=1}^{\infty} \mathbf{A}_j \mathbf{X}_j$ *and* $L_2 = \sum_{j=1}^{\infty} \mathbf{B}_j \mathbf{X}_j$ *where* $\mathbf{A}_j, \mathbf{B}_j$ *are non-singular* $d \times d$ *matrices. If* L_1 *and* L_2 *are independent and* $\{\mathbf{A}_j \mathbf{B}_j^{-1}\}_{j \geq 1}$ *as well as* $\{\mathbf{B}_j \mathbf{A}_j^{-1}\}_{j \geq 1}$ *are bounded in some matrix norm, then the random vectors* $\mathbf{X}_1, \mathbf{X}_2, \ldots$ *are normally distributed.*

B. Proofs

B.1. Proof of Lemma 1.2

Proof. The proof follows its univariate version (Peters et al., 2009). We first show that if $\mathbf{X}_t = \sum_{j=0}^{\infty} \mathbf{\Psi}_j \mathbf{Z}_{t-j}$, then $\mathbf{Z}_t$ is independent of $\mathbf{X}_i$, for all $i < t$. By defining $\mathbf{X}_t^{(n)} := \sum_{j=0}^{n} \mathbf{\Psi}_j \mathbf{Z}_{t-j}$, we have that $\mathbf{X}_t^{(n)}$ converges weakly to $\mathbf{X}_t$ and thus the characteristic function of $(\mathbf{X}_t, \mathbf{Z}_{t+1})$ obeys for all u and v

$$\begin{aligned}
\varphi_{\mathbf{P}^{(\mathbf{X}_t, \mathbf{Z}_{t+1})}}(u,v) &= \lim_{n\to\infty} \varphi_{\mathbf{P}^{(\mathbf{X}_t^{(n)}, \mathbf{Z}_{t+1})}}(u,v) \\
&= \lim_{n\to\infty} \varphi_{\mathbf{P}^{\mathbf{X}_t^{(n)}}}(u) \varphi_{\mathbf{P}^{\mathbf{Z}_{t+1}}}(v) \\
&= \varphi_{\mathbf{P}^{\mathbf{X}_t}}(u) \varphi_{\mathbf{P}^{\mathbf{Z}_{t+1}}}(v) \\
&= \varphi_{\mathbf{P}^{\mathbf{X}_t} \otimes \mathbf{P}^{\mathbf{Z}_{t+1}}}(u,v).
\end{aligned}$$

This results in independence of $\mathbf{Z}_t$ and $X_i, \; i < t$ by uniqueness of the characteristic function.

In order to prove the "if"-part, we need to show that for a causal process $\mathbf{X}_t$, all coefficients $\mathbf{\Psi}_i$ are equal to zero for $i < 0$ in the Laurent expansion $\mathbf{X}_t = \sum_{i \in \mathbb{Z}} \mathbf{\Psi}_i \mathbf{Z}_{t-i}$ (see (Lütkepohl, 2010)). Assume otherwise, i.e., there is a coefficient $i_0 < 0$ such that $\mathbf{\Psi}_{i_0} \neq 0$. Then

$$\mathbf{\Psi}_{i_0} \mathbf{Z}_{t-i_0} + \sum_{i \in \mathbb{Z} - i_0} \mathbf{\Psi}_i \mathbf{Z}_{t-i} = \mathbf{X}_t \perp\!\!\!\perp \mathbf{\Psi}_{i_0} \mathbf{Z}_{t-i_0}. \quad (10)$$

Because $\mathbf{\Psi}_{i_0} \mathbf{Z}_{t-i_0}$ and $\sum_{i \in \mathbb{Z} - i_0} \mathbf{\Psi}_i \mathbf{Z}_{t-i}$ are independent with the same reasoning as above, (10) results in a contradiction. □

B.2. Proof of Proposition 2.1

Proof. Defining

$$\widetilde{\mathbf{Z}}_t := \mathbf{X}_t - \mathrm{cov}(\mathbf{X}_t, \mathbf{X}_{t+1}) \cdot \mathrm{cov}(\mathbf{X}_{t+1}, \mathbf{X}_{t+1})^{-1} \mathbf{X}_{t+1}$$

and

$$\widetilde{\mathbf{\Phi}} := \mathrm{cov}(\mathbf{X}_t, \mathbf{X}_{t+1}) \cdot \mathrm{cov}(\mathbf{X}_{t+1}, \mathbf{X}_{t+1})^{-1},$$

it follows that

$$\begin{aligned}
\widetilde{\mathbf{\Phi}} \mathbf{X}_{t+1} + \widetilde{\mathbf{Z}}_t &= \mathrm{cov}(\mathbf{X}_t, \mathbf{X}_{t+1}) \cdot \mathrm{cov}(\mathbf{X}_{t+1}, \mathbf{X}_{t+1})^{-1} \mathbf{X}_{t+1} \\
&\quad + \mathbf{X}_t - \mathrm{cov}(\mathbf{X}_t, \mathbf{X}_{t+1}) \mathrm{cov}(\mathbf{X}_{t+1}, \mathbf{X}_{t+1})^{-1} \mathbf{X}_{t+1} = \mathbf{X}_t.
\end{aligned}$$

In addition, we have

$$\begin{aligned}
\mathrm{cov}(\widetilde{\mathbf{Z}}_t, \mathbf{X}_{t+1}) &= \mathrm{cov}(\mathbf{X}_t - \mathrm{cov}(\mathbf{X}_t, \mathbf{X}_{t+1}) \\
&\quad \cdot \mathrm{cov}(\mathbf{X}_{t+1}, \mathbf{X}_{t+1})^{-1} \mathbf{X}_{t+1}, \mathbf{X}_{t+1}) \\
&= \mathrm{cov}(\mathbf{X}_t, \mathbf{X}_{t+1}) - \mathrm{cov}(\mathbf{X}_t, \mathbf{X}_{t+1}) \\
&\quad \cdot \mathrm{cov}(\mathbf{X}_{t+1}, \mathbf{X}_{t+1})^{-1} \mathrm{cov}(\mathbf{X}_{t+1}, \mathbf{X}_{t+1}) \\
&= 0.
\end{aligned}$$

By the assumption of the Gaussian distribution, the independence of $\widetilde{\mathbf{Z}}_t$ and $\mathbf{X}_{t+1}$ follows. It remains to show that $\widetilde{\mathbf{Z}}_t$ and $\mathbf{X}_{t+k}$ are independent for $k \geq 2$. By the multivariate form of the Yule-Walker equations for the $\mathrm{VAR}(1)$ process, i.e., $\mathbf{\Gamma}_k := \mathrm{cov}(\mathbf{X}_t, \mathbf{X}_{t+k}) = \mathbf{\Phi} \mathbf{\Gamma}_{k-1}$ (see 2.1.31 in (Lütkepohl, 2010)), we obtain that

$$\begin{aligned}
\mathrm{cov}(\widetilde{\mathbf{Z}}_t, \mathbf{X}_{t+k}) &= \mathrm{cov}(\mathbf{X}_t - \mathrm{cov}(\mathbf{X}_t, \mathbf{X}_{t+1}) \\
&\quad \cdot \mathrm{cov}(\mathbf{X}_{t+1}, \mathbf{X}_{t+1})^{-1} \mathbf{X}_{t+1}, \mathbf{X}_{t+k}) \\
&= \mathrm{cov}(\mathbf{X}_t, \mathbf{X}_{t+k}) - \mathrm{cov}(\mathbf{X}_t, \mathbf{X}_{t+1}) \\
&\quad \cdot \mathrm{cov}(\mathbf{X}_{t+1}, \mathbf{X}_{t+1})^{-1} \mathrm{cov}(\mathbf{X}_{t+1}, \mathbf{X}_{t+k}) \\
&= \mathrm{cov}(\mathbf{X}_t, \mathbf{X}_{t+k}) - \mathbf{\Phi} \mathbf{\Gamma}_0 \mathbf{\Gamma}_0^{-1} \\
&\quad \cdot \mathrm{cov}(\mathbf{X}_{t+1}, \mathbf{X}_{t+k}) = \mathbf{\Gamma}_k - \mathbf{\Phi} \mathbf{\Gamma}_{k-1} = 0.
\end{aligned}$$

□

B.3. Proof of Lemma 3.1

Proof. The statement follows from considering determinants, see Lemma 1.1.

$$\begin{aligned}
\det(\mathbb{1}_{K(p+q)} - \mathbf{\Upsilon} z) &= \det(\mathbb{1}_{Kp} - \mathbf{\Upsilon}_{11} z) \\
&\quad \cdot \det(\mathbb{1}_{Kq} - \mathbf{\Upsilon}_{22} z) \\
&= \det(\mathbb{1}_K - \mathbf{\Phi}_1 z - \cdots - \mathbf{\Phi}_p z^p)
\end{aligned}$$

□

C. Additional Formula to Section 3.2

The following submatrices of $\boldsymbol{\Upsilon} = \begin{bmatrix} \boldsymbol{\Upsilon}_{11} & \boldsymbol{\Upsilon}_{12} \\ \mathbf{0}_{Kq \times Kp} & \boldsymbol{\Upsilon}_{22} \end{bmatrix}$ are used to represent a VARMA process as a VAR process of order one:

$$\boldsymbol{\Upsilon}_{11} := \begin{bmatrix} \boldsymbol{\Phi}_1 & \boldsymbol{\Phi}_2 & \dots & \boldsymbol{\Phi}_{p-1} & \boldsymbol{\Phi}_p \\ \mathbb{1}_K & 0 & \dots & 0 & 0 \\ 0 & \mathbb{1}_K & \dots & 0 & 0 \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \dots & 0 & \mathbb{1}_K & 0 \end{bmatrix},$$

$$\boldsymbol{\Upsilon}_{12} := \begin{bmatrix} \boldsymbol{\Theta}_1 & \boldsymbol{\Theta}_2 & \dots & \boldsymbol{\Theta}_{q-1} & \boldsymbol{\Theta}_q \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

and

$$\boldsymbol{\Upsilon}_{22} := \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ \mathbb{1}_K & 0 & \dots & 0 & 0 \\ 0 & \mathbb{1}_K & \dots & 0 & 0 \\ \vdots & & \ddots & 0 & \vdots \\ 0 & \dots & 0 & \mathbb{1}_K & 0 \end{bmatrix}.$$