Statistical Tests for the Detection of the Arrow of Time in Vector Autoregressive Models

Pablo Morales-Mombiela, Daniel Hernández-Lobato and Alberto Suárez

Escuela Politécnica Superior, Universidad Autónoma de Madrid, Spain

Causal Discovery in the Context of Multi-dimensional Time-series

Given a sample of a stationary multi-variate time series

$$\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N, \quad (1)$$

we analyze if this sequence is in the correct chronological order or its time ordering has been reversed. This is a particular case of the general causal inference problem. In the one-dimensional case it has been shown that under a stationary AR model of the form:

$$X_t = \phi X_{t-1} + \epsilon_t, \quad \epsilon_t \perp X_{t-1}, \quad (2)$$

with non-Gaussian i.i.d. noise ϵ_t, the residuals in the backward time direction

$$\tilde{\epsilon}_t \equiv X_t - \phi X_{t+1}, \quad t = 1, 2, \ldots, T, \quad (3)$$

are more Gaussian than the corresponding residuals in the forward direction, $\{\epsilon_t\}_{t=1}^{T}$ [Hernández-Lobato *et al.*, 2011]. In particular, the magnitude of the cumulants of order higher than $\mathbf{2}$ is reduced:

$$\kappa_n(\tilde{\epsilon}_t) = c_n(\phi)\kappa_n(\epsilon_t), \quad n > 0$$
$$c_n(\phi) = (-\phi)^n + (1-\phi^2)^n(1-\phi^n)^{-1}, \quad (4)$$

where $\kappa_n(\cdot)$ denotes the n-th cumulant. For a stationary AR(1) processes with $\phi \neq 0$ and $|\phi| < 1$ this implies that

$$|\kappa_n(\tilde{\epsilon}_t)| \leq |\kappa_n(\epsilon_t)|, \quad \forall n > 2. \quad (5)$$

Time Reversal in Vector Autoregressive Models

Consider a d-dimensional autoregressive model

$$\mathbf{X}_t = \mathbf{A}\mathbf{X}_{t-1} + \epsilon_t, \quad \epsilon_t \perp \mathbf{X}_{t-1}, \quad (6)$$

where $\mathbf{A}$ is a matrix whose eigenvalues are within the unit circle in the complex plane and ϵ_t is i.i.d noise. A linear fit of the time-reversed process is:

$$\mathbf{X}_t = \tilde{\mathbf{A}}\mathbf{X}_{t+1} + \tilde{\epsilon}_t, \quad (7)$$

with $\tilde{\epsilon}_t$ time-reversed residuals. The matrix of autoregressive coefficients for the reversed time series, $\tilde{\mathbf{A}}$, needs not be equal to $\mathbf{A}$. Both matrices are related by:

$$\tilde{\mathbf{A}} = \Sigma\mathbf{A}'\Sigma^{-1}, \quad \Sigma = \mathbb{E}[\mathbf{X}_t\mathbf{X}_t']. \quad (8)$$

One can define the time-reversed residuals of a linear fit

$$\tilde{\epsilon}_t \equiv \mathbf{X}_t - \tilde{\mathbf{A}}\mathbf{X}_{t+1}. \quad (9)$$

- $\tilde{\epsilon}_t$ is Gaussian and $\tilde{\epsilon}_t \perp \mathbf{X}_{t+1}$ if and only if ϵ_t is multi-dimensional Gaussian i.i.d noise.
- Otherwise, $\tilde{\epsilon}_t$ is not Gaussian i.i.d noise and $\tilde{\epsilon}_t \not\perp \mathbf{X}_{t+1}$.

By analogy to the one-dimensional case we conjecture that the (multi-dimensional) distribution of backward residuals $\{\tilde{\epsilon}_t\}$ is more Gaussian than the distribution of forward residuals $\{\epsilon_t\}$ and propose to use this to determine the direction of time.

Statistical Tests for the Detection of the Arrow of Time

Tests Based on Independence: The residuals in the each direction satisfy $\epsilon_t \perp \mathbf{X}_{t-1}$ and $\tilde{\epsilon}_t \not\perp \mathbf{X}_{t+1}$. Thus, tests of independence, *e.g.* the HSIC [Gretton *et al.*, 2008], can be used to determine the correct direction of time [Peters *et al.*, 2009].

Tests Based on Measures of Gaussianity: The test performs a linear fit in the original and in the reversed ordering. The ordering in which the residuals are less Gaussian is then chosen. A measure of deviation from the Gaussian distribution is used.

A measure of Deviation from a Bi-variate Gaussian Distribution

Theorem: Let ϵ^x and ϵ^y be two random variables. Let $\mathbf{Z}(\alpha) \equiv \epsilon^x \cos\alpha + \epsilon^y \sin\alpha$. $\mathbf{Z}(\alpha)$ is normal $\forall\alpha \in [0, \pi]$ if and only if the joint distribution of ϵ^x and ϵ^y is normal.

Inspired by this theorem we use the integrated excess of kurtosis to estimate the deviation from the bi-variate Gaussian of the distribution of the residuals $\epsilon_t = (\epsilon_t^x, \epsilon_t^y)$:

$$\int |\kappa_4| \equiv \frac{1}{\pi}\int_0^\pi d\alpha |\kappa_4[\mathbf{Z}(\alpha)]|. \quad (10)$$

The computational cost is $\mathcal{O}(N)$, *i.e.*, linear in the number of observations.

Experimental Settings: Noise and Auto-correlation matrix A

$\mathbf{A} = \mathbf{PDP}^{-1}$ with $\mathbf{D} = \mathrm{diag}(\lambda_1, \lambda_2)$. λ_1, λ_2 and $\mathbf{P}$ are fixed to specific values.

Two types of noise are considered:

$$\epsilon_t = \begin{pmatrix} \epsilon_t^x \\ \epsilon_t^y \end{pmatrix} = \begin{pmatrix} \mathrm{sign}(Z_t^x)|Z_t^x|^{r_x} \\ \mathrm{sign}(Z_t^y)|Z_t^y|^{r_y} \end{pmatrix},$$

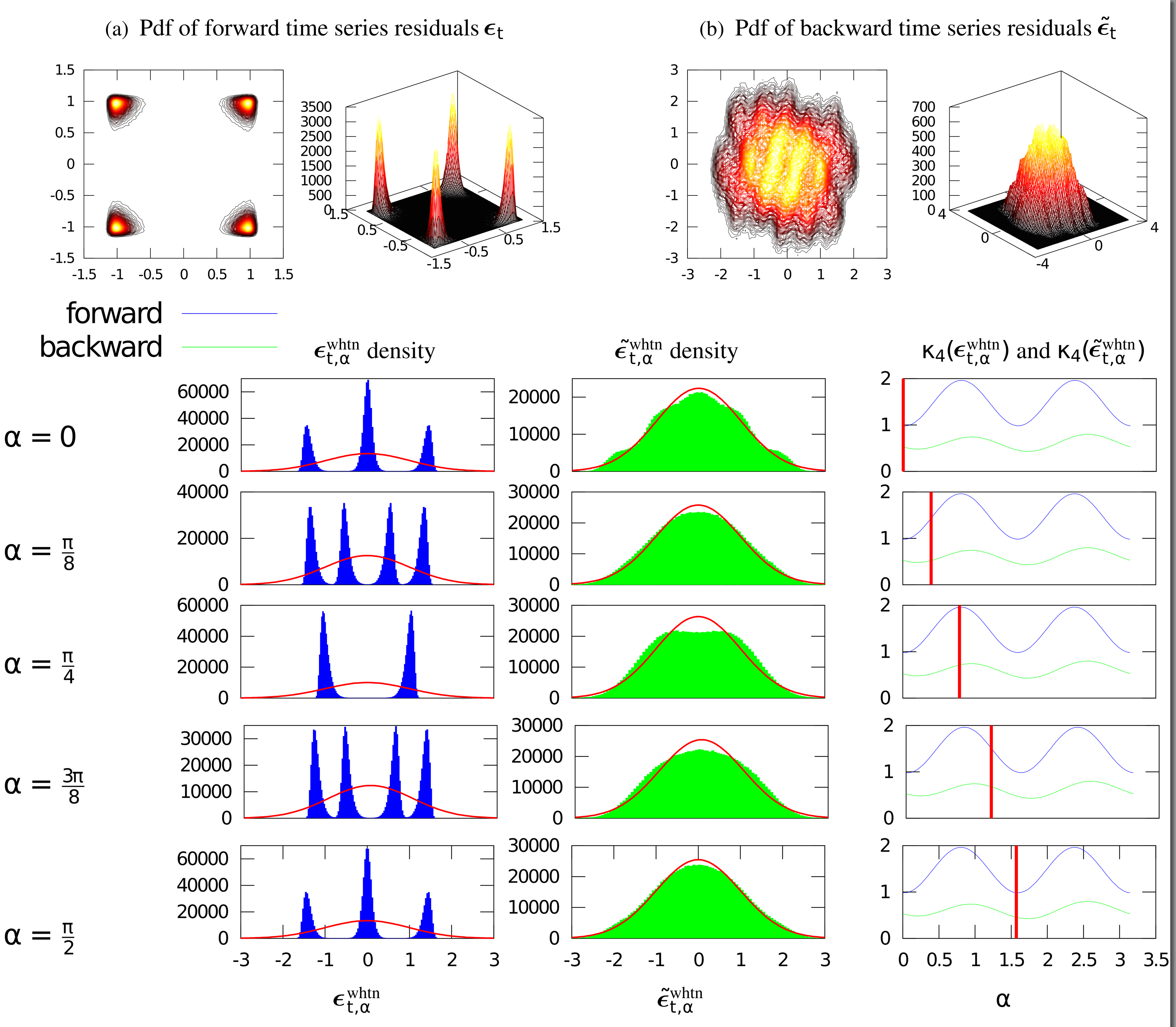
where $(Z_t^x, Z_t^y)^\top \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$

$$\epsilon_t \sim C(\Phi(\epsilon_t^x), \Phi(\epsilon_t^y); \theta),$$

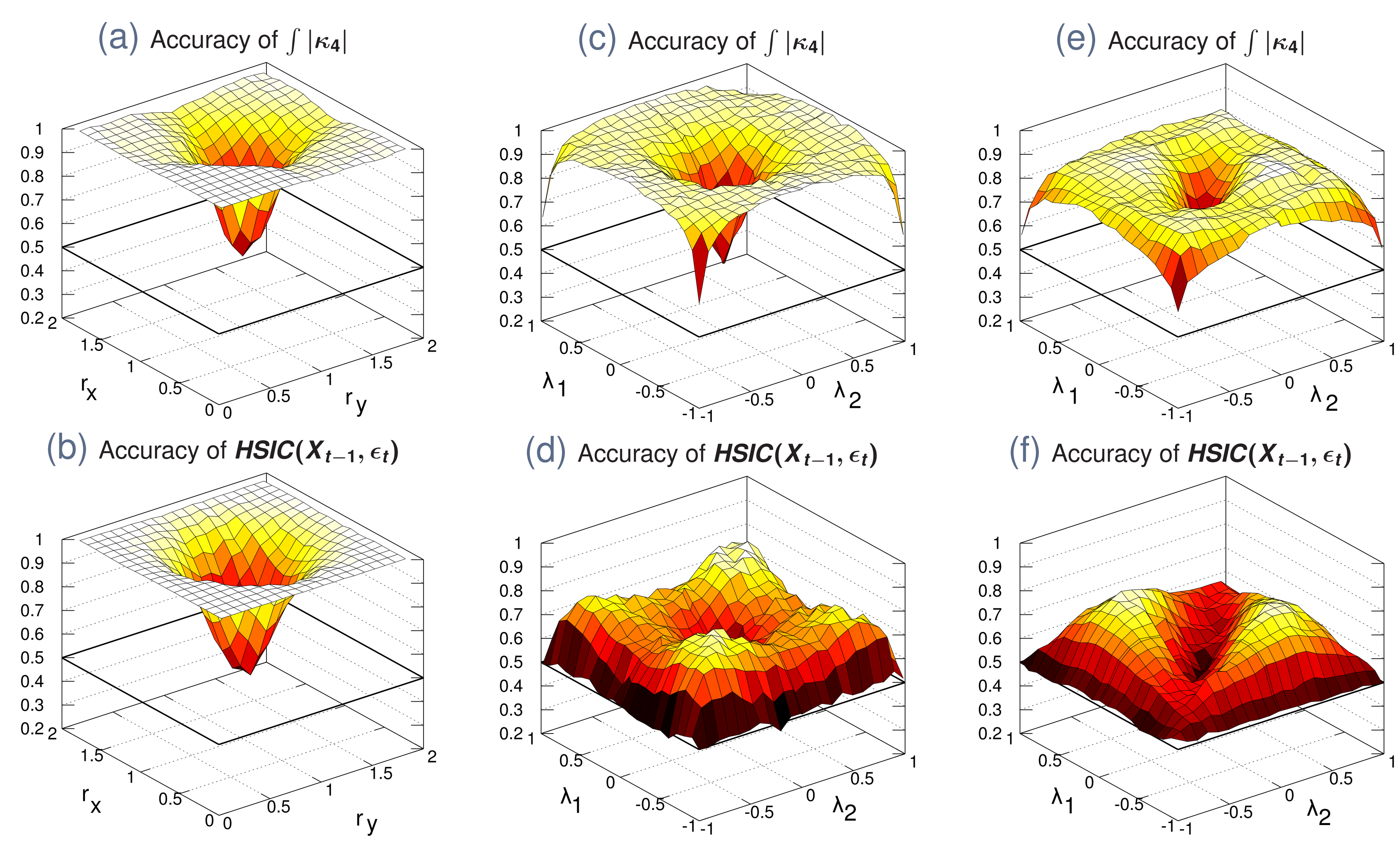
where $(\epsilon_t^x, \epsilon_t^y)^\top \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. $C(\cdot, \cdot; \theta)$ denotes a Frank copula with parameter θ and $\Phi(\cdot)$ is the cpf of a normal Gaussian.

r_x or r_y determine the level of non-Gaussianity of the marginals, from fully Gaussian ($r_x = r_y = 1$) to leptokurtic ($r_x > 1$ and $r_y > 1$) or platokurtic ($r_x < 1$ and $r_y < 1$). The second type of noise has Gaussian marginals but non-Gaussian dependencies.

Gaussianization Effect in the Time-reversed Residuals



Synthetic Experiments: Results



In experiments (a) and (b) $\mathbf{P} = \mathbf{I}$ and $\lambda_1 = \lambda_2 = (\sqrt{5} - 1)/2$. In experiments (c) and (d) $r_x = r_y = 0.75$ and $\mathbf{P} \neq \mathbf{I}$. In these experiments the first type of noise is employed. In experiments (e) and (f) $\mathbf{P} = \mathbf{I}$ and the second type of noise is employed with $\theta = 10$.

Conclusions

- A statistical test determines the direction of time of a multi-variate time series generated by a VAR$_d$(1). The direction of time is the one in which the residuals are less Gaussian.
- A measure of discrepancy between the distribution of the residuals from a multi-variate Gaussian distribution has been defined.
- Tests based on measures of Gaussianity show better performance than tests based on measures of independence. Furthermore, they are more efficient.
- If $\mathbf{X}$ and $\mathbf{Y}$ are identically distributed random variables and the relation between them is linear, the the proposed test can be used to determine whether $\mathbf{X}$ causes $\mathbf{Y}$.

References

- J.M. Hernández-Lobato, P. Morales-Mombiela, and A. Suárez. Gaussianity measures for detecting the direction of causal time series. In 22nd IJCAI, 2011.
- Peters, D. Janzing, A. Gretton, and B. Schölkopf. Detecting the direction of causal time series. In ICML , pages 801-808. ACM, 2009.
- A. Gretton, K. Fukumizu, C. H. Teo, L. Song, B. Schölkopf, and A. Smola. A kernel statistical test of independence. In NIPS 20, pages 585-592. MIT Press, 2008.

http://www.eps.uam.es/%7ega/ pablo.morales@estudiante.uam.es