SORT 40 (2) July-December 2016, 267-302

A construction of continuous-time ARMA models by iterations of Ornstein-Uhlenbeck processes

Argimiro Arratia[1], Alejandra Cabaña[2] and Enrique M. Cabaña[3]

Abstract

We present a construction of a family of continuous-time ARMA processes based on p iterations of the linear operator that maps a Lévy process onto an Ornstein-Uhlenbeck process. The construction resembles the procedure to build an AR(p) from an AR(1). We show that this family is in fact a subfamily of the well-known CARMA(p,q) processes, with several interesting advantages, including a smaller number of parameters. The resulting processes are linear combinations of Ornstein-Uhlenbeck processes all driven by the same Lévy process. This provides a straightforward computation of covariances, a state-space model representation and methods for estimating parameters. Furthermore, the discrete and equally spaced sampling of the process turns to be an ARMA($p,p-1$) process. We propose methods for estimating the parameters of the iterated Ornstein-Uhlenbeck process when the noise is either driven by a Wiener or a more general Lévy process, and show simulations and applications to real data.

MSC: 60G10, 62M10, 62M99 60M99.

Keywords: Ornstein-Uhlenbeck process, Lévy process, Continuous ARMA, stationary process.

1. Introduction

The link between discrete time autoregressive moving average (ARMA) processes and stationary processes with continuous-time has been of interest for many years, see for instance, Doob (1944), Durbin (1961), Bergstrom (1984, 1996) and more recently Brockwell (2009), Thornton and Chambers (2013). Continuous time ARMA processes are better suited than their discrete counterparts for modelling irregularly spaced data, and when the white noise is driven by a non-Gaussian process it becomes a more realistic model in finance and other fields of application.

[1] Universitat Politècnica de Catalunya, Barcelona, Spain. argimiro@cs.upc.edu
Supported by Spain's MINECO project APCOM (TIN2014-57226-P) and Generalitat de Catalunya 2014SGR 890 (MACDA).

[2] Universitat Autònoma de Barcelona, Spain. acabana@mat.uab.cat
Supported by Spain's MINECO project MTM2015-69493-R.

[3] Universidad de la República, Montevideo, Uruguay. ecabana@ccee.edu.uy
Received: November 2015
Accepted: May 2016

A popular continuous-time representation of ARMA(p,q) process (known as CARMA(p,q)) can be obtained via a state-space representation of the formal equation

$$a(D)Y(t) = \sigma b(D)D\Lambda(t),$$

where $\sigma > 0$ is a scale parameter, D denotes differentiation with respect to t, Λ is a second-order Lévy process, $a(z) = z^p + a_1 z^{p-1} + \ldots + a_p$ is a polynomial of order p and $b(z) = b_0 + b_1 z + \ldots + b_q z^q$ a polynomial of order $q \leq p-1$ with coefficient $b_q \neq 0$ (see, e.g., Brockwell, 2004, 2009, Thornton and Chambers, 2013). The parameters of this model are estimated by adjusting first an ARMA(p,q), $q < p$ to regularly spaced data. Then obtain the parameters of the continuous version whose values at the observation times have the same distribution of the fitted ARMA. Hence, $p+q+1$ parameters have to be estimated.

We propose in this work a parsimonious model for continuous autoregression, with fewer parameters (as we shall see exactly p plus the variance). Our construction departs from the observation that a Ornstein-Uhlenbeck (OU) process can be thought of as continuous-time interpolation of an autoregressive process of order one (i.e. an AR(1)). This is shown in Section 2, where we also review some well known facts on Lévy processes, ARMA models and their representations. The model is obtained by a procedure that resembles the one that allows to build an AR(p) from an AR(1). Departing from this analogy, we define and analyse the result of iterating the application of the operator that maps a Wiener process onto an OU process. This operator is defined in Section 3 and denoted $\mathcal{OU}$, with subscripts denoting the parameters involved.

The p iterations of $\mathcal{OU}$, for each positive integer p, give rise to a new family of processes, the Ornstein-Uhlenbeck processes of order p, denoted OU(p). They can be used as models for either stationary continuous-time processes or the series obtained by observing these continuous processes at equally spaced instants. We show that an OU(p) process can be expressed as a linear combination of ordinary OU processes, or generalized OU processes, also defined in Section 3. This result resembles the aggregations of Gaussian (and non-Gaussian) processes studied with the idea of deconstructing a complicated economic model into simpler constituents. In the extensive literature on aggregations (or superpositions) of stochastic processes the aggregated processes are driven by independent Lévy processes (see, e.g., Granger and Morris, 1976, Granger, 1980, Barndorff-Nielsen, 2001, Eliazar and Klafter, 2009, among many others). A distinctive point of our construction is that the stochastic processes obtained by convolution of the OU operator result in a linear combination comprised of processes driven by the same Lévy process.

Another consequence of writing the OU(p) process as the aggregation of simpler ones is the derivation of a closed formula for its covariance. This has important practical implications since it allows to easily estimate the parameters of a OU(p) process by matching correlations (a procedure resembling the method of moments, to be described in Section 6.2), and by maximum likelihood.

In Section 4 we show how to write the discrete version of a OU(p) as a state-space model, and from this representation we show in Section 5 that for $p > 1$, a OU(p) behaves like an aggregation of AR processes (in the manner considered in Granger and Morris (1976)), that turns out to be an ARMA(p,q), with $q \leq p-1$. Consequently the OU(p) processes are a subfamily of the CARMA(p,q) processes. Notwithstanding this structural similarity, the family of discretized OU(p) processes is more parsimonious than the family of ARMA$(p,p-1)$ processes, and we shall see empirically that it is able to fit well the autocovariances for large lags. Hence, OU processes of higher order appear as a new continuous model, competitive in a discrete time setting with higher order autoregressive processes (AR or ARMA). The estimation of the parameters of OU(p) processes is attempted in Section 6. Simulations and applications to real data are provided in Section 6.5. Our concluding remarks are in Section 7.

2. Preliminaries

Let us recall that a Lévy process $\Lambda(t)$ is a càdlàg function, with independent and stationary increments, that vanishes in $t = 0$. As a consequence, $\Lambda(t)$ is, for each t, a random variable with an infinitely divisible law (Sato, 1999). A Wiener process W is a centred Gaussian process, with independent increments and variance $\mathrm{E}(W(t) - W(s))^2 = \sigma^2|t-s|$. Wiener processes are the only Lévy processes with almost surely continuous paths. For parameter $\lambda > 0$ the classical Ornstein-Uhlenbeck process is defined as $\int_{-\infty}^{t} \mathrm{e}^{-\lambda(t-s)} dW(s)$ (Uhlenbeck and Ornstein, 1930).

Wiener process can be replaced by a second order Lévy process Λ to define a Lévy driven Ornstein-Uhlenbeck process as

$$x(t)(= x_{\lambda,\Lambda}(t)) := \int_{-\infty}^{t} \mathrm{e}^{-\lambda(t-s)} d\Lambda(s) \tag{1}$$

The previous equation can be formally written in differential form

$$dx(t) = -\lambda x(t)dt + d\Lambda(t) \tag{2}$$

We may think of x as the result of accumulating a random noise $d\Lambda$, with reversion to the mean (that we assume to be 0) of exponential decay with rate λ.

When the Ornstein-Uhlenbeck process x is sampled at equally spaced times $\{h\tau : h = 0,1,2,\ldots,n\}$, $\tau > 0$, the series $X_h = x(h\tau)$ obeys an autoregressive model of order 1 (i.e. an AR(1)), because $X_{h+1} = e^{-\lambda\tau} X_h + Z_{h+1}$, where $Z_{h+1} = \int_{h\tau}^{(h+1)\tau} e^{-\lambda((h+1)\tau - s)} d\Lambda(s)$, is the stochastic innovation.

Hence, we can consider the OU process as continuous-time interpolation of an AR(1) process. Notice that both models are stationary. This link between AR(1) and OU(1) suggests the definition of iterated OU processes introduced in Section 3.

An ARMA(p,q) or autoregressive moving average process of order (p,q) has the following form

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \theta_0 \epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}$$

where $\phi_1, \ldots, \phi_p$ are the autoregressive parameters, $\theta_0, \ldots, \theta_q$ are the moving average parameters, and the white-noise process ϵ_t has variance one. Denote by B the backshift operator that carries x_t into x_{t-1}. By considering the polynomials in the backshift operator,

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p \text{ and } \theta(B) = \theta_0 + \theta_1 B + \cdots + \theta_q B^q$$

the ARMA(p,q) model can be written as

$$\phi(B) x_t = \theta(B) \epsilon_t \tag{3}$$

This compact expression comes in handy for analysing structural properties of time series. It also links to the representation of ARMA processes as a *state-space model*, useful for simplifying maximum likelihood estimation and forecasting. A state-space model has the general form

$$\boldsymbol{Y}_t = \boldsymbol{A}\boldsymbol{Y}_{t-1} + \boldsymbol{\eta}_t \tag{4}$$

$$x_t = \boldsymbol{K}^\top \boldsymbol{Y}_t + \boldsymbol{N}_t \tag{5}$$

where (4) is the state equation and (5) is the observation equation, with $\boldsymbol{Y}_t$ the m-dimensional state vector, $\boldsymbol{A}$ and $\boldsymbol{K}$ are $m \times m$ and $m \times k$ coefficient matrices, $\boldsymbol{K}^\top$ denotes the transpose of $\boldsymbol{K}$, $\boldsymbol{\eta}$ and $\boldsymbol{N}$ are m and k dimensional white noises. $\boldsymbol{N}$ would be present only if the process x_t is observed subject to additional noise (see Box, Jenkins, and Reinsel, 1994 for further details). We present in Section 4 a state-space model representation of our generalized OU process.

3. Ornstein-Uhlenbeck processes of order p

The AR(1) process $X_t = \phi X_{t-1} + \epsilon_t$, where ϵ_t, $t \in$Z, is a white noise, can be written as $(1 - \phi B)X_t = \epsilon_t$ using the back-shift operator B. Equivalently, X_t can be written as $X_t = \mathcal{MA}_{1/\rho}\epsilon_t$, where $\mathcal{MA}_{1/\rho}$ is the moving average that maps ϵ_t onto $\mathcal{MA}_{1/\rho}\epsilon_t, = \sum_{j=0}^{\infty} \frac{1}{\rho^j}\epsilon_{t-j}$, and ρ $(= 1/\phi)$ is the root of the characteristic polynomial $1 - \phi z$.

Moreover, the AR(p) process $X_t = \sum_{j=1}^{p} \phi_j X_{t-j} + \epsilon_t$ (or $\phi(B)X_t = \epsilon_t$), where $\phi(z) = 1 - \sum_{j=1}^{p} \phi_j z^j = \prod_{j=1}^{p}(1 - z/\rho_j)$ has roots $\rho_j = \mathrm{e}^{\lambda_j}$, $j = 1, \ldots, p$, can be obtained by applying the composition of the moving averages $\mathcal{MA}_{1/\rho_j}$ to the noise, that is:

$$X_t = \prod_{j=1}^{p} \mathcal{MA}_{1/\rho_j} \epsilon_t$$

Now consider the operator $\mathcal{MA}_{\mathrm{e}^{-\lambda}}$ that maps ϵ_t onto

$$\mathcal{MA}_{\mathrm{e}^{-\lambda}} \epsilon_t = \sum_{l \le t, \text{integer}} \mathrm{e}^{-\lambda(t-l)} \epsilon_l$$

A continuous version of this operator is $\mathcal{OU}_\lambda$ that maps $y(t), t \in \mathbb{R}$ onto

$$\mathcal{OU}_\lambda y(t) = \int_{-\infty}^{t} \mathrm{e}^{-\lambda(t-s)} dy(s), \tag{6}$$

whenever the integral can be defined. The definition of $\mathcal{OU}_\lambda$ is extended to include complex processes, by replacing λ by $\kappa = \lambda + i\mu$, $\lambda > 0$, $\mu \in \mathbb{R}$ in (6). The set of complex numbers with positive real part is denoted by $\mathbf{C}^+$, and the conjugate of κ is denoted by $\bar{\kappa}$.

For $p \ge 1$ and parameters $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_p)$, the previous argument suggests to define the following process obtained as repeated compositions of operators $\mathcal{OU}_{\kappa_j}$, $j = 1, \ldots, p$:

$$\mathcal{OU}_{\boldsymbol{\kappa}} y(t) := \mathcal{OU}_{\kappa_1} \mathcal{OU}_{\kappa_2} \cdots \mathcal{OU}_{\kappa_p} y(t) = \prod_{j=1}^{p} \mathcal{OU}_{\kappa_j} y(t) \tag{7}$$

This is called *Ornstein-Uhlenbeck process of order p with parameters* $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_p) \in (\boldsymbol{C}^+)^p$. The composition $\prod_{j=1}^{p} \mathcal{OU}_{\kappa_j}$ is unambiguously defined because the application of $\mathcal{OU}_{\kappa_j}$ operators is commutative as shown in Theorem 1(i) below.

The particular case of interest where the underlying noise is a second order Lévy process Λ, namely,

$$\mathcal{OU}_{\boldsymbol{\kappa}} \Lambda(t) := \mathcal{OU}_{\kappa_1} \mathcal{OU}_{\kappa_2} \cdots \mathcal{OU}_{\kappa_p} \Lambda(t) = \prod_{j=1}^{p} \mathcal{OU}_{\kappa_j} \Lambda(t) \tag{8}$$

is called the *Lévy-driven Ornstein-Uhlenbeck process of order p with parameters* $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_p) \in (\boldsymbol{C}^+)^p$.

For technical reasons, it is convenient to introduce the *Ornstein-Uhlenbeck operator* $\mathcal{OU}_\kappa^{(h)}$ *of degree h with parameter* κ that maps y onto

$$\mathcal{OU}_\kappa^{(h)} y(t) = \int_{-\infty}^{t} \mathrm{e}^{-\kappa(t-s)} \frac{(-\kappa(t-s))^h}{h!} dy(s) \tag{9}$$

and Λ onto

$$\xi_\kappa^{(h)}(t) = \int_{-\infty}^{t} \mathrm{e}^{-\kappa(t-s)} \frac{(-\kappa(t-s))^h}{h!} d\Lambda(s) \tag{10}$$

We call the process (10) *generalized Ornstein-Uhlenbeck process* of order 1 and degree h. For the remainder of the paper we restrict the underlying noise to a second order Lévy Λ, but note that the general properties of the $\mathcal{OU}_\kappa$ operator that we are going to show hold for any random function $y(t)$ for which the integral (6) is defined.

3.1. Properties of the operator $\mathcal{OU}_\kappa$

The following statements summarize some properties of products (compositions) of the operators defined by (7) and (9), and correspondingly, of the stationary centred processes $\xi_\kappa^{(h)}$, $h \geq 0$. In particular, the Ornstein-Uhlenbeck processes of order 1 and degree 0, $\xi_\kappa^{(0)} = \xi_\kappa$ are the ordinary Ornstein-Uhlenbeck processes (1).

Theorem 1

(i) When $\kappa_1 \neq \kappa_2$*, the product* $\mathcal{OU}_{\kappa_2}\mathcal{OU}_{\kappa_1}$ *can be computed as*

$$\frac{\kappa_1}{\kappa_1 - \kappa_2}\mathcal{OU}_{\kappa_1} + \frac{\kappa_2}{\kappa_2 - \kappa_1}\mathcal{OU}_{\kappa_2}$$

and is therefore commutative.

(ii) The composition $\prod_{j=1}^{p} \mathcal{OU}_{\kappa_j}$ *constructed with pairwise different* $\kappa_1, \ldots, \kappa_p$ *is equal to the linear combination*

$$\prod_{j=1}^{p} \mathcal{OU}_{\kappa_j} = \sum_{j=1}^{p} K_j(\kappa_1, \ldots, \kappa_p)\mathcal{OU}_{\kappa_j}, \tag{11}$$

with coefficients

$$K_j(\kappa_1, \ldots, \kappa_p) = \frac{1}{\prod_{\kappa_l \neq \kappa_j}(1 - \kappa_l/\kappa_j)}. \tag{12}$$

(iii) *For* $i = 1,2,\ldots$, $\mathcal{OU}_\kappa \mathcal{OU}_\kappa^{(i)} = \mathcal{OU}_\kappa^{(i)} - \kappa \mathcal{OU}_\kappa^{(i+1)}$.

(iv) *For any positive integer* p *the* p*-th power of the Ornstein-Uhlenbeck operator has the expansion*

$$\mathcal{OU}_\kappa^p = \sum_{j=0}^{p-1} \binom{p-1}{j} \mathcal{OU}_\kappa^{(j)}. \tag{13}$$

(v) *Let* $\kappa_1, \ldots, \kappa_q$ *be pairwise different complex numbers with positive real parts, and* $p_1, \ldots, p_q$ *positive integers, and let us denote by* $\boldsymbol{\kappa}$ *a complex vector in* $(\boldsymbol{C}^+)^p$ *with components* κ_h *repeated* p_h *times,* $p_h \geq 1$, $h = 1, \ldots, q$, $\sum_{h=1}^q p_h = p$. *Then, with* $K_h(\boldsymbol{\kappa})$ *defined by (12),*

$$\prod_{h=1}^q \mathcal{OU}_{\kappa_h}^{p_h} = \sum_{h=1}^q \frac{1}{\prod_{l \neq h}(1 - \kappa_l/\kappa_h)^{p_l}} \mathcal{OU}_{\kappa_h}^{p_h} = \sum_{h=1}^q K_h(\boldsymbol{\kappa}) \mathcal{OU}_{\kappa_h}^{p_h}.$$

An immediate consequence is that the operator $\mathcal{OU}_{\boldsymbol{\kappa}}$ with p-vector parameter $\boldsymbol{\kappa}$ can be written as a linear combination of p operators $\mathcal{OU}_\kappa$ or $\mathcal{OU}_\kappa^{(h)}$ for suitable scalar values κ and non-negative integer h. Therefore, the process $\mathcal{OU}_{\boldsymbol{\kappa}}\Lambda$ can be written as a linear combination of OU processes driven by the same Lévy process, as stated in the following Corollary.

Corollary 1

(i) *The process* $\mathcal{OU}_{\boldsymbol{\kappa}}(\Lambda) = \prod_{h=1}^q \mathcal{OU}_{\kappa_h}^{p_h}(\Lambda)$ *can be expressed as the linear combination*

$$\mathcal{OU}_{\boldsymbol{\kappa}}(\Lambda) = \sum_{h=1}^q K_h(\boldsymbol{\kappa}) \sum_{j=0}^{p_h-1} \binom{p_h-1}{j} \xi_{\kappa_h}^{(j)} \tag{14}$$

of the p *processes* $\{\xi_{\kappa_h}^{(j)} : h = 1, \ldots, q, j = 0 \ldots, p_h - 1\}$ *(see (10)).*

(ii) *Consequently,*

$$\mathcal{OU}_{\boldsymbol{\kappa}}\Lambda(t) = \sum_{h=1}^q K_h(\boldsymbol{\kappa}) \sum_{j=0}^{p_h-1} \binom{p_h-1}{j} \int_{-\infty}^t \mathrm{e}^{-\kappa_h(t-s)} \frac{(-\kappa_h(t-s))^j}{j!} d\Lambda(s)$$

Corollary 2 *For real* λ, μ, *with* $\lambda > 0$, *the product* $\mathcal{OU}_{\lambda+i\mu}\mathcal{OU}_{\lambda-i\mu}$ *is real, that is, applied to a real process produces a real image.*

The proofs of Theorem 1 and corollaries are in Appendix A.

3.2. Computing the covariances

The representation

$$x := \mathcal{OU}_{\boldsymbol{\kappa}}(\Lambda) = \sum_{h=1}^{q} K_h(\kappa) \sum_{j=0}^{p_h-1} \binom{p_h-1}{j} \mathcal{OU}_{\kappa_h}^{(j)}(\Lambda)$$

of x as a linear combination of the processes $\xi_{\kappa_h}^{(i)} = \mathcal{OU}_{\kappa_h}^{(i)}(\Lambda)$ allows a direct computation of the covariances $\gamma(t) = \mathrm{E}x(t)\bar{x}(0)$ through a closed formula, in terms of the covariances $\gamma_{\kappa_1,\kappa_2}^{(i_1,i_2)}(t) = \mathrm{E}\xi_{\kappa_1}^{(i_1)}(t)\bar{\xi}_{\kappa_2}^{(i_2)}(0)$:

$$\gamma(t) = \sum_{h'=1}^{q} \sum_{i'=0}^{p_{h'}-1} \sum_{h''=1}^{q} \sum_{i''=0}^{p_{h''}-1} K_{h'}(\boldsymbol{\kappa}) \bar{K}_{h''}(\boldsymbol{\kappa}) \binom{p_{h'}-1}{i'} \binom{p_{h''}-1}{i''} \gamma_{\kappa_{h'},\kappa_{h''}}^{(i',i'')}(t) \tag{15}$$

with $v^2 = \mathrm{Var}\Lambda(1)$,

$$\begin{aligned}
\gamma_{\kappa_1,\kappa_2}^{(i_1,i_2)}(t) &= v^2(-\kappa_1)^{i_1}(-\bar{\kappa}_2)^{i_2} \int_{-\infty}^{0} \mathrm{e}^{-\kappa_1(t-s)} \frac{(t-s)^{i_1}}{i_1!} \mathrm{e}^{-\bar{\kappa}_2(-s)} \frac{(-s)^{i_2}}{i_2!} ds \\
&= v^2(-\kappa_1)^{i_1}(-\bar{\kappa}_2)^{i_2} \mathrm{e}^{-\kappa_1 t} \sum_{j=0}^{i_1} \binom{i_1}{j} \frac{t^j}{i_1! i_2!} \int_{-\infty}^{0} \mathrm{e}^{(\kappa_1+\bar{\kappa}_2)s} (-s)^{i_1+i_2-j} ds \\
&= \frac{v^2(-\kappa_1)^{i_1}(-\bar{\kappa}_2)^{i_2} \mathrm{e}^{-\kappa_1 t}}{i_2!} \sum_{j=0}^{i_1} \frac{t^j (i_1+i_2-j)!}{j!(i_1-j)!(\kappa_1+\bar{\kappa}_2)^{(i_1+i_2-j+1)}}
\end{aligned} \tag{16}$$

A real expression for the covariance when the imaginary parameters appear as conjugate pairs can be obtained but it is much more involved than this one.

4. The OU(p) process as a state-space model

Theorem 1 and its corollaries lead to express the OU(p) process by means of linear state-space models. The state-space modelling provides a unified methodology for the analysis of time series (see Durbin and Koopman, 2001).

In the simplest case, where the elements of $\boldsymbol{\kappa}$ are all different, the process $x(t) = \mathcal{OU}_{\boldsymbol{\kappa}}\Lambda(t)$ is a linear combination of the state vector $\boldsymbol{\xi}_{\boldsymbol{\kappa}}(t) = (\xi_{\kappa_1}(t), \xi_{\kappa_2}(t), \ldots, \xi_{\kappa_p}(t))^{\intercal}$, where $\xi_{\kappa_j} = \mathcal{OU}_{\kappa_j}(\Lambda)$.

More precisely, the vectorial process

$$\boldsymbol{\xi}_{\boldsymbol{\kappa}}(t) = (\xi_{\kappa_1}(t), \xi_{\kappa_2}(t), \ldots, \xi_{\kappa_p}(t))^{\intercal},\ \xi_{\kappa_j} = \mathcal{OU}_{\kappa_j}(\Lambda)$$

and $x(t) = \mathcal{OU}_{\boldsymbol{\kappa}}\Lambda(t)$ satisfy the linear equations

$$\boldsymbol{\xi}_{\boldsymbol{\kappa}}(t) = \operatorname{diag}(\mathrm{e}^{-\kappa_1\tau}, \mathrm{e}^{-\kappa_2\tau}, \ldots, \mathrm{e}^{-\kappa_p\tau})\boldsymbol{\xi}_{\boldsymbol{\kappa}}(t-\tau) + \boldsymbol{\eta}_{\boldsymbol{\kappa},\boldsymbol{\tau}}(t) \tag{17}$$

$$\text{and} \qquad x(t) = \boldsymbol{K}^{\intercal}(\boldsymbol{\kappa})\boldsymbol{\xi}(t), \tag{18}$$

$$\boldsymbol{\eta}_{\boldsymbol{\kappa},\boldsymbol{\tau}}(t) = (\eta_{\kappa_1,\tau}(t), \eta_{\kappa_2,\tau}(t), \ldots, \eta_{\kappa_p,\tau}(t))^{\intercal},\ \eta_{\kappa_j,\tau}(t) = \int_{t-\tau}^{t} \mathrm{e}^{-\kappa_j(t-s)} d\Lambda(s),$$

$$\operatorname{Var}(\boldsymbol{\eta}_{\boldsymbol{\kappa},\boldsymbol{\tau}}(t)) = v^2((v_{j,l})),\ v_{j,l} = \int_{t-\tau}^{t} \mathrm{e}^{-(\kappa_j+\bar{\kappa}_l)(t-s)} ds = \frac{1-\mathrm{e}^{-(\kappa_j+\bar{\kappa}_l)\tau}}{\kappa_j+\bar{\kappa}_l} \tag{19}$$

and the coefficients from (12), $\boldsymbol{K}^{\intercal}(\boldsymbol{\kappa}) = (K_1(\boldsymbol{\kappa}), K_2(\boldsymbol{\kappa}), \ldots, K_p(\boldsymbol{\kappa}))$.

The initial value $\boldsymbol{\xi}(0)$ is estimated by means of its conditional expectation $\hat{\boldsymbol{\xi}}(0) = \mathrm{E}(\xi(0)|x(0)) = \frac{\boldsymbol{K}^{\intercal}(\boldsymbol{\kappa})Vx(0)}{\boldsymbol{K}^{\intercal}(\boldsymbol{\kappa})V\overline{\boldsymbol{K}}}$, with $V = \operatorname{Var}(\boldsymbol{\xi}(0)) = \left(\left(\frac{1}{\kappa_j+\bar{\kappa}_l}\right)\right)$.

An application of Kalman filter to this state-space model leads to compute the likelihood of $\boldsymbol{x} = (x(0), x(\tau), \ldots, x(n\tau))$. Some Kalman filter programs included in software packages require the processes in the state-space to be real. That condition is not fulfilled by the model described by equations (17) and (18). An equivalent description by means of real processes can be obtained by ordering the parameters $\boldsymbol{\kappa}$ with the imaginary components paired with their conjugates in such a way that $\kappa_{2h} = \bar{\kappa}_{2h-1}$, $h = 1, 2, \ldots, c$ and the imaginary component $\Im(\kappa_j) = 0$ if and only if $2c < j \leq p$.

Then the matrix $\boldsymbol{M} = ((M_{j,k}))$ with all elements equal to zero except $M_{2h-1,2h-1} = M_{2h-1,2h} = 1$, $-M_{2h,2h-1} = M_{2h,2h} = i$, $h = 1, 2, \ldots, c$ and $M_{j,j} = 1$, $2c < j \leq p$, induces the linear transformation $\boldsymbol{\xi} \mapsto \boldsymbol{M\xi}$ that leads to the new state-space description

$$\boldsymbol{M\xi}(t) = \boldsymbol{M}\operatorname{diag}(\mathrm{e}^{-\kappa_1\tau}, \mathrm{e}^{-\kappa_2\tau}, \ldots, \mathrm{e}^{-\kappa_p\tau})\boldsymbol{M}^{-1}\boldsymbol{M\xi}(t-\tau) + \boldsymbol{M\eta}(t), \tag{20}$$

$$x(t) = \boldsymbol{K}^{\intercal}\boldsymbol{M}^{-1}M\boldsymbol{\xi}(t), \tag{21}$$

where the processes $\boldsymbol{M\xi}$ are real.

Observe that there is no loss of generality in choosing the spacing τ between observations as unity for the derivation of the state-space equations. Hence, we set $\tau = 1$ in the sequel and, in addition, τ will be omitted from the notation.

When $\kappa_1, \ldots, \kappa_q$ are all different, $p_1, \ldots, p_q$ are positive integers, $\sum_{h=1}^{q} p_h = p$ and $\boldsymbol{\kappa}$ is a p-vector with p_h repeated components equal to κ_h, the OU(p) process $x(t) = \mathcal{OU}_{\boldsymbol{\kappa}}\Lambda(t)$ is a linear function of the state-space vector

$$\left(\xi_{\kappa_1}^{(0)}, \xi_{\kappa_1}^{(1)}, \ldots, \xi_{\kappa_1}^{(p_1-1)}, \ldots, \xi_{\kappa_q}^{(0)}, \xi_{\kappa_q}^{(1)}, \ldots, \xi_{\kappa_q}^{(p_q-1)}\right)$$

where the components are given by (10), and the transition equation is no longer expressed by a diagonal matrix. In this case the state-space model has the following form

$$\begin{aligned} \boldsymbol{\xi}(t) &= \boldsymbol{A}\boldsymbol{\xi}(t-1) + \boldsymbol{\eta}(t) \\ x(t) &= \boldsymbol{K}^{\intercal}\boldsymbol{\xi}(t) \end{aligned} \tag{22}$$

We leave the technical details of this derivation to Appendix B. The terms $\boldsymbol{\xi}(t)$, $\boldsymbol{A}$, $\boldsymbol{\eta}(t)$ and $\boldsymbol{K}$ are precisely defined in (36). The real version of (22), when the process $\boldsymbol{\xi}$ has imaginary components is obtained by multiplying both equations by a block-diagonal matrix $\boldsymbol{C}$ (which is defined precisely in the Appendix), giving us the real state-space model

$$\boldsymbol{C}\boldsymbol{\xi}(t) = (\boldsymbol{C}\boldsymbol{A}\boldsymbol{C}^{-1})(\boldsymbol{C}\boldsymbol{\xi}(t-1)) + \boldsymbol{C}\boldsymbol{\eta}(t), \tag{23}$$

$$x(t) = (\boldsymbol{K}^{\intercal}C^{-1})(C\boldsymbol{\xi}(t)). \tag{24}$$

5. The OU(p) as an ARMA(p, $p-1$)

The studies of properties of linear transformations and aggregations of similar processes have produced a great amount of work stemming from the seminal paper by Granger and Morris (1976) on the invariance of MA and ARMA processes under these operations. These results and extensions to vector autoregressive moving average (VARMA) processes are compiled in the textbook by Lütkepohl (2005).

The description of the OU(p) process $x = \mathcal{OU}_{\boldsymbol{\kappa}}(\Lambda)$ with parameters $\boldsymbol{\kappa}$ as a linear state-space model, given in the previous section, will allow us to show that the series $x(0)$, $x(1)$, …, $x(n)$ satisfies an ARMA(p,q) model with q smaller than p. We refer the reader to (Lütkepohl, 2005, Ch. 11) for a presentation on VARMA processes and, in particular, to the following result on the invariance property of VARMA processes under linear transformations, which we quote with a minor change of notation:

Theorem 2 (Lütkepohl, 2005, Cor. 11.1.2) *Let $y(t)$ be a d-dimensional, stable, invertible VARMA($\tilde{p}$,$\tilde{q}$) process and let $\boldsymbol{F}$ be an $(m \times d)$ matrix of rank m. Then the process $z_t = \boldsymbol{F}y_t$ has a VARMA($\check{p},\check{q}$) representation with $\check{p} \leq (d-m+1)\tilde{p}$ and $\check{q} \leq (d-m)\tilde{p}+\tilde{q}$.* ■

Equation (23) shows that $C\boldsymbol{\xi}(t)$ is a p-dimensional autoregressive vector (a p-dimensional VARMA(1,0) process) and Equation (24) expresses $x(t)$ as a linear transformation of $C\boldsymbol{\xi}(t)$ by the $(1 \times p)$ matrix $\boldsymbol{F} = \boldsymbol{K}^\intercal \boldsymbol{C}^{-1}$. Using Theorem 2 (with $d = p$, $\tilde{p} = 1$, $\tilde{q} = 0$, $m = 1$) we conclude that $(x(t) : t = 0, 1, \ldots, n)$ is an ARMA($\check{p}$,$\check{q}$) process with $\check{p} \leq p$ and $\check{q} \leq p-1$:

$$x(h) = \sum_{j=1}^{p} \phi_j x(h-j) + \sum_{l=0}^{p-1} \theta_l \epsilon_{h-l} \tag{25}$$

where ϵ is a Gaussian white noise with variance 1 and the parameters $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_p)^\intercal$, $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_{p-1})^\intercal$ of the ARMA process are functions of the parameters $\boldsymbol{\kappa}$ of the OU process. When the noise is any other second order Lévy process the corresponding OU(p) process has the same covariances as the process (25).

By using the backshift operator B, and the polynomials $\phi(z) = 1 - \sum_{j=1}^{p} \phi_j z^j$, $\theta(z) = \sum_{l=0}^{p-1} \theta_l z^l$, (25) is written as

$$\phi(B)x = \theta(B)\epsilon. \tag{26}$$

5.1. Identifying the ARMA(p, $p-1$) from a given OU(p) process

We proceed now to identify the coefficients $\boldsymbol{\phi} \in \mathbb{R}^p$ and $\boldsymbol{\theta} \in \mathbb{R}^{p-1}$ of the ARMA$(p, p-1)$ model that has the same autocovariances as $x = \mathcal{OU}_{\boldsymbol{\kappa}}(\Lambda)$.

Case 1. Consider first that all components of $\boldsymbol{\kappa}$ are pairwise different, and hence $x(t)$ $= \sum_{j=1}^{p} K_j \xi_{\kappa_j}(t)$ is a linear combination of the OU(1) processes

$$\xi_{\kappa_j}(t) = \int_{-\infty}^{t} \mathrm{e}^{-\kappa_j(t-s)} d\Lambda(s) = \mathrm{e}^{-\kappa_j} \xi_{\kappa_j}(t-1) + \int_{t-1}^{t} \mathrm{e}^{-\kappa_j(t-s)} d\Lambda(s)$$

with innovations $\boldsymbol{\eta}_{\boldsymbol{\kappa}}$ with components $\eta_{\kappa_j}(t) = \int_{t-1}^{t} \mathrm{e}^{-\kappa_j(t-s)} d\Lambda(s)$.

For each j, the series $\xi_{\kappa_j} = (\xi_{\kappa_j}(h))_{h \in \boldsymbol{Z}}$ satisfies the AR(1) model

$$(1 - \mathrm{e}^{-\kappa_j} B)\xi_{\kappa_j} = \eta_{\kappa_j}$$

(see (17)), and from (18) the series $x = (x(h))_{h\in\mathbf{Z}}$ follows the ARMA model

$$\prod_{j=1}^{p}(1-\mathrm{e}^{-\kappa_j}B)x = \sum_{j=1}^{p} K_j(\boldsymbol{\kappa})\prod_{l\neq j}(1-\mathrm{e}^{-\kappa_l}B)\eta_{\kappa_j}.$$

The sum of moving averages in the right-hand term is distributed as the moving average

$$\zeta = \sum_{h=0}^{p-1}\theta_h B^h\epsilon$$

where ϵ is a white noise with variance one and the coefficients θ_h are suitably chosen. It is readily verified that the autocovariances $c_l = \mathrm{E}\zeta(h)\bar{\zeta}(h-l)$ of this MA are the coefficients in the sum of powers of z

$$\left(\sum_{h=0}^{p-1}\theta_h z^h\right)\left(\sum_{k=0}^{p-1}\bar{\theta}_k z^{-h}\right) = \sum_{l=-p+1}^{p-1} c_l z^l. \tag{27}$$

A similar formula that takes into account the correlations (19) between the noises η_{κ_k} indicates that the same autocovariances are given by the identity

$$J(z) := \sum_{j=1}^{p}\sum_{l=1}^{p} K_j\bar{K}_l G_j(z)\bar{G}_l(1/z)v_{j,l} = \sum_{l=-p+1}^{p-1} c_l z^l \tag{28}$$

where $G_j(z) = \prod_{l\neq j}(1-\mathrm{e}^{-\kappa_l}z) = \sum_{l=0}^{p-1} g_{j,l}z^l$.

The coefficients $g_{j,l}$ and the function J are completely determined from the parameters of the OU process. In order to express the parameters of the ARMA$(p,p-1)$ process in terms of $\boldsymbol{\kappa}$ and $v^2 = \mathrm{Var}\Lambda(1)$ it remains to obtain the coefficients θ_h in the factorization (27). The roots ρ_j $(j = 1,2,\ldots,p-1)$ of

$$\theta(z) = \sum_{h=0}^{p-1}\theta_h z^h = \theta_0\prod_{j=1}^{p-1}(1-z/\rho_j) \tag{29}$$

are obtained by choosing the roots of the polynomial $z^{p-1}\theta(z)\bar{\theta}(1/z) = z^{p-1}J(z)$ with modules greater than one (the remaining roots are their inverses). Then all θ_h are written in terms of the ρ_h and the size factor θ_0 by applying (29). The value of θ_0 follows by using an additional equation, namely, the equality of the terms of degree zero in $J(z)$ and $\theta(z)\bar{\theta}(1/z)$, thus obtaining

$$\sum_{l=0}^{p-1}|\theta_l|^2=\sum_{j=1}^{p}\sum_{l=1}^{p}K_j\bar{K}_l v_{j,l}\sum_{h=0}^{p-1}g_{j,h}\bar{g}_{l,h}.$$

The general result, for arbitrary $\boldsymbol{\kappa}$ is much more involved and its derivation is deferred to Appendix C.

6. Estimation of the parameters of OU(p)

6.1. Reparameterization by means of real parameters

Our purpose is to insert the expression (15) for the covariance $\gamma(t)$ of the process $x(t)=\mathcal{OU}_{\boldsymbol{\kappa}}\Lambda(t)$ in a numeric optimization procedure in order to compute the maximum likelihood estimates of the parameters. Although $\gamma(t)$ depends continuously on $\boldsymbol{\kappa}$, the same does not happen with each term in the expression (15), because of the lack of boundedness of the coefficients of the linear combination when two different values of the components of $\boldsymbol{\kappa}$ approach each other. Since we wish to consider real processes x and the process itself and its covariance $\gamma(t)$ depend only of the unordered set of the components of $\boldsymbol{\kappa}$, we shall reparameterize the process. For the sake of simplicity, but without losing generality, consider the case where the components in $\boldsymbol{\kappa}$ are pairwise different. Let $K_{j,i}=\frac{1}{(-\kappa_j)^i\prod_{l\neq j}(1-\kappa_l/\kappa_j)}$ (in particular, $K_{j,0}$ is the same as K_j). Then the processes $x_i(t)=\sum_{j=1}^{p}K_{j,i}\xi_j(t)$ and the coefficients $\boldsymbol{\beta}=(\beta_1,\ldots,\beta_p)$ of the polynomial

$$g(z)=\prod_{j=1}^{p}(1+\kappa_j z)=1-\sum_{j=1}^{p}\beta_j z^j. \tag{30}$$

satisfy

$$\sum_{i=1}^{p}\beta_i x_i(t)=x(t).$$

The resulting process is real, because of Corollary 2. This works likewise for the general case of $\boldsymbol{\kappa}$ with some repetitions. Therefore the new parameter $\boldsymbol{\beta}$ shall be adopted.

6.2. Matching correlations estimation (MCE)

From the closed formula (15) for the covariance γ and the relationship (30) between $\boldsymbol{\kappa}$ and $\boldsymbol{\beta}$, we have a mapping $(\boldsymbol{\beta},v^2)\mapsto\gamma(t)$, for each t. Since

$$\boldsymbol{\rho}^{(T)}:=(\rho(1),\rho(2),\ldots,\rho(T))^{\intercal}=(\gamma(1),\gamma(2),\ldots,\gamma(T))^{\intercal}/\gamma(0)$$

does not depend on v^2, these equations determine a map $\mathcal{C} : (\boldsymbol{\beta}, T) \mapsto \boldsymbol{\rho}^{(T)} = \mathcal{C}(\boldsymbol{\beta}, T)$ for each T. After choosing a value of T and obtaining an estimate $\boldsymbol{\rho}_e^{(T)}$ of $\boldsymbol{\rho}^{(T)}$ based on the empirical covariances of x, we propose as a first estimate of $\boldsymbol{\beta}$, the vector $\breve{\boldsymbol{\beta}}_T$ such that all the components of the corresponding $\boldsymbol{\kappa}$ have positive real parts, and such that the Euclidean norm $\|\boldsymbol{\rho}_e^{(T)} - \mathcal{C}(\breve{\boldsymbol{\beta}}_T, T)\|$ reaches its minimum. The procedure resembles the estimation by the *method of moments*. The components of $\boldsymbol{\rho}_e^{(T)}$ for the series $(x_j)_{j=1,2,\ldots,n}$ are computed as

$$\rho_{e,h} = \gamma_{e,h}/\gamma_{e,0}, \qquad \gamma_{e,h} = \frac{1}{n}\sum_{j=1}^{n-h} x_j x_{j+h}.$$

6.3. Maximum likelihood estimation (MLE) in the Gaussian case

In this case $x(t) = \mathcal{OU}_{\boldsymbol{\kappa}} \sigma W(t)$, where $W(t)$ is standard Wiener process. Assume that $x(t)$ is observed at times $0, \tau, 2\tau, \ldots, n\tau$. By choosing τ the time unit of measure, as in Section 4, we assume without loss of generality that our observations are $\boldsymbol{x} = (x(0), x(1), \ldots, x(n))^\intercal$.

The likelihood L of the vector $\boldsymbol{x}$ is given by

$$\log L(\boldsymbol{x}; \boldsymbol{\beta}, \sigma^2) = -\tfrac{n}{2}\log(2\pi) - \tfrac{1}{2}\log(\det(\boldsymbol{\Gamma}(\boldsymbol{\beta}, \sigma^2)) - \tfrac{1}{2}\boldsymbol{x}^\intercal(\boldsymbol{\Gamma}(\boldsymbol{\beta}, \sigma^2))^{-1}\boldsymbol{x}$$

where $\boldsymbol{\Gamma}$ has components $\Gamma_{h,j} = \gamma(|h-j|)$ $(h, j = 0, 1, \ldots, n)$. The Kalman filter associated to the dinamical state-space model in Section 4 provides an efficient alternative to compute the likelihood.

From these elements, a numerical optimization leads to obtain the maximum likelihood estimators $\hat{\boldsymbol{\beta}}$ of $\boldsymbol{\beta}$ and $\hat{\sigma}^2$ of σ^2. If required, the estimations $\hat{\boldsymbol{\kappa}}$ follow by solving the analogue of the polynomial equation (30) written in terms of the estimators:

$$\prod_{j=1}^{p}(1 + \hat{\kappa}_j z) = 1 - \sum_{j=1}^{p}\hat{\beta}_j z^j.$$

The optimization for large n and the solution of the algebraic equation for large p require a considerable computation effort, but there are efficient programs to perform both operations, such as `optim` and `polyroot` in `R` (R Core Team, 2015). An alternative when the observed process is not assumed to be centred, is to maximize the log-likelihood of $\Delta\boldsymbol{x} = (x(1) - x(0), x(2) - x(1), \ldots, x(n) - x(n-1))$ given by

$$\log L(\boldsymbol{x}; \boldsymbol{\beta}, \sigma^2) = -\tfrac{n}{2}\log(2\pi) - \tfrac{1}{2}\log(\det(\boldsymbol{V}(\boldsymbol{\beta}, \sigma^2)) - \tfrac{1}{2}\Delta\boldsymbol{x}^\intercal(\boldsymbol{V}(\boldsymbol{\beta}, \sigma^2))^{-1}\Delta\boldsymbol{x}$$

with $\boldsymbol{V}(\boldsymbol{\beta},\sigma^2)$ equal to the $n\times n$ matrix with components

$$V_{h,j} = 2\gamma(|h-j|) - \gamma(|h-j|+1) - \gamma(|h-j|-1)$$

that reduce to $2(\gamma(0)-\gamma(1))$ at the diagonal $h=j$.

The optimization procedures require an initial guess about the value of the parameter to be estimated. The estimators obtained by matching correlations described in the previous section can be used for that purpose.

6.4. The Gaussian case: examples

When Λ is a Wiener process W, the OU process of order p belongs to a subclass with $p+1$ parameters of the classical family of the $2p$-parameters Gaussian ARMA$(p,p-1)$

$$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \theta_0\epsilon_t + \theta_1\epsilon_{t-1} + \cdots + \theta_{p-1}\epsilon_{t-p+1}$$

where $\phi_1,\ldots,\phi_p$ and $\theta_0,\ldots,\theta_q$ are parameters and ϵ_t is a Gaussian noise with variance 1. The parameters $\boldsymbol{\kappa},\sigma^2$ determine the Gaussian likelihood of $\mathcal{OU}_{\boldsymbol{\kappa}}\sigma W$, and are estimated by the values $\hat{\boldsymbol{\kappa}}$ and $\hat{\sigma}^2$ that maximize that likelihood.

We have observed in several examples that the covariances of the process with the maximum likelihood estimators as parameters, follow closely the empirical covariances of the series. We have simulated the sample paths for the Wiener-driven OU(p) for different values of the parameters.

In the examples below we present simulated series $x(j), j=0,1,2,\ldots,n$ obtained from an OU process x for $n=300$ and three different values of the parameters and computed the MC and ML estimators $\check{\boldsymbol{\beta}}_T$, and $\hat{\boldsymbol{\beta}}$. The value of T for the MC estimation has been arbitrarily set equal to the integral part of $0.9\cdot n$, but the graphs of $\check{\boldsymbol{\beta}}_T$ for several values of T show in each case that after T exceeds a moderate threshold, the estimates remain practically constant. One of such graphs is included below (see Figure 2). It is of interest to perform further comparisons of these two methodologies for parameter estimation. A recent antecedent of this kind of comparisons and its importance can be found in Nieto, Orbe and Zarraga (2014).

The simulations show that the correlations of the series with the estimated parameters are fairly adapted to each other and to the empirical covariances. The departure from the theoretical covariances of x can be ascribed to the simulation intrinsic randomness.

Our first two examples describe OU(3) processes with arbitrarily (and randomly) chosen parameters and the third one imitates the behaviour of Series A that appears in Section 6.5.

Example 1. A series $(x_h)_{h=0,1,\dots,n}$ of $n = 300$ observations of the $\mathrm{OU}_{\boldsymbol{\kappa}}$ process x of order $p = 3$, $\boldsymbol{\kappa} = (0.9, 0.2+0.4i, 0.2-0.4i)$ and $\sigma^2 = 1$ was simulated, and the parameters $\boldsymbol{\beta} = (-1.30, -0.56, -0.18)$ and $\sigma^2 = 1$ were estimated by means of matching correlations:

$$\check{\boldsymbol{\beta}}_T = (-1.9245, -0.6678, -0.3221),$$

with $T = 270$; and maximum likelihood:

$$\hat{\boldsymbol{\beta}} = (-1.3546, -0.6707, -0.2355)$$

and $\hat{\sigma}^2 = 0.8958$. The corresponding estimators for $\boldsymbol{\kappa}$ are $\check{\boldsymbol{\kappa}} = (1.6368, 0.1439 + 0.4196i, 0.14389 - 0.4196i)$ and $\hat{\boldsymbol{\kappa}} = (0.9001, 0.2273 + 0.4582i, 0.2273 - 0.4582i)$.

The following table summarizes the different estimations of this OU(3) process.

original $\boldsymbol{\beta}$	-1.30	-0.56	-0.18	$\sigma^2 = 1$
original $\boldsymbol{\kappa}$	0.9	$0.2+0.4i$	$0.2-0.4i$	$\sigma^2 = 1$
MCE $\check{\boldsymbol{\beta}}_T$	-1.9245	-0.6678	-0.3221	
$\check{\boldsymbol{\kappa}}$	1.6368	$0.1439+0.4196i$	$0.14389-0.4196i$	
MLE $\hat{\boldsymbol{\beta}}$	-1.3546	-0.6707	-0.2355	$\hat{\sigma}^2 =$
$\hat{\boldsymbol{\kappa}}$	0.9001	$0.2273+0.4582i$	$0.2273-0.4582i$	0.8958

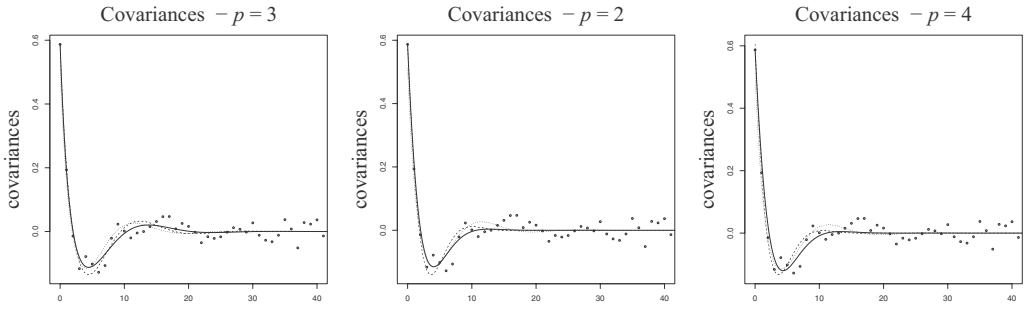


Figure 1: *Empirical covariances (∘) and covariances of the MC (—) and ML (- - -) fitted OU models, for $p = 3$, 2 and 4 corresponding to Example 1. The covariances of $\mathrm{OU}_{\boldsymbol{\kappa}}$ are indicated with a dotted line.*

Figure 1 describes the theoretical, empirical and estimated covariances of x under the assumption $p = 3$, that is, the actual order of x. The results obtained when the estimation is performed for $p = 2$ and $p = 4$ are also shown. Figure 2 shows that the MC estimates of $\boldsymbol{\beta}$ become stable for T moderately large, and close to the already indicated estimations for $T = 270$ (the horizontal lines).

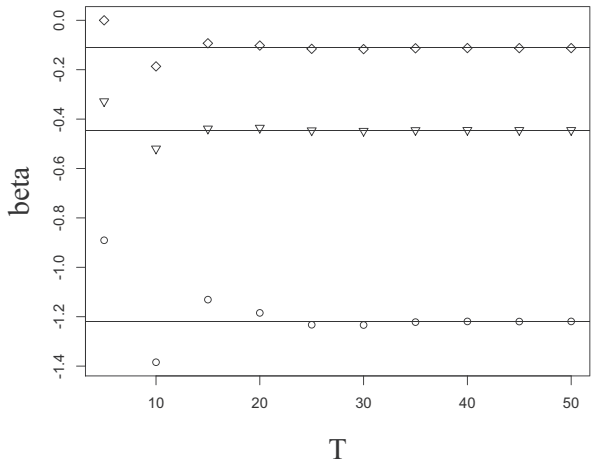


Figure 2: *The MC estimations $\check{\beta}_1(\circ)$, $\check{\beta}_2(\triangledown)$ and $\check{\beta}_3(\diamond)$ for different values of T, corresponding to Example 1. The horizontal lines indicate the estimations for $T = 270$.*

The coefficients ϕ_1, ϕ_2, ϕ_3 of the ARMA(3,2) model (26) satisfied by the series $(x(h))_{h=0,1,\ldots,300}$ are obtained by computing the product $\prod_{j=1}^{3}(1-\mathrm{e}^{-\kappa_j}B) = 1-\phi_1 B - \phi_2 B^2 - \phi_3 B^3 = 1 - 1.9148B + 1.2835B^2 - 0.2725B^3$.

As for the coefficients $\theta_0, \theta_1, \theta_2$, the first step is to compute the function

$$J(z) = 0.2995z^{-2} - 1.1943z^{-1} + 1.7904 - 1.1943z + 0.2995z^2,$$

then obtain the roots $\rho_1 = 1.1443 - 0.1944i$, $\rho_2 = 1.1443 + 0.1944i$, $\rho_3 = 0.8494 - 0.1443i$, $\rho_4 = 0.8494 + 0.1443i$ of the equation $z^2 J(z) = 0$, ordered by decreasing moduli, discard the last two, and write the function $\theta(z) = \theta_0 + \theta_1 z + \theta_2 z^2$ defined in (29):

$$\theta_0 \prod_{j=1}^{2}(1 - B/\rho_j) = \theta_0(1 - 1.6988z + 0.7423z^2).$$

Solve $\theta_0^2(1 + (-1.6988)^2 + 0.74229^2) = 1.7904$ to have $\theta_0 = 0.6352$, and hence $\theta(B) = 0.6352 - 1.0791B + 0.4715B^2$.

Example 2. The process $x = \mathrm{OU}_{(0.04,0.21,1.87)}$ is analysed as in Example 1. The resulting estimators are $\check{\boldsymbol{\beta}}_T = (-2.0611, -0.7459, -0.0553)$, $T = 270$, $\check{\boldsymbol{\kappa}} = (1.6224, 0.3378, 0.1009)$, $\hat{\boldsymbol{\beta}} = (-1.8253, -0.7340, -0.0680)$, $\hat{\sigma}^2 = 0.7842$, $\hat{\boldsymbol{\kappa}} = (1.3015, 0.3897, 0.1342)$.

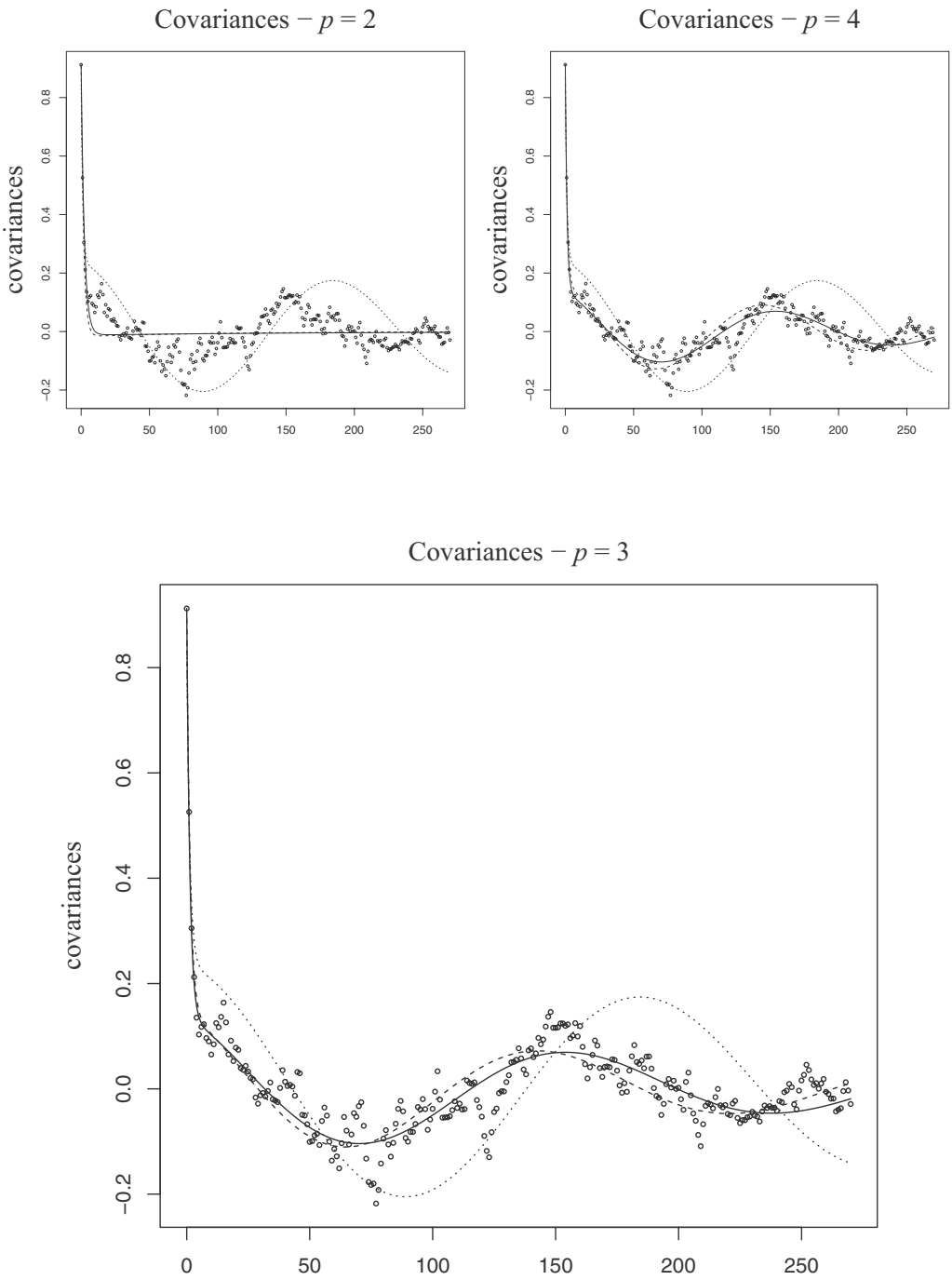


Figure 3: *Empirical covariances (○) and covariances of the MC (—) and ML (- - -) fitted OU models, for $p = 2, p = 4$ and $p = 3$, the actual value of the parameter, corresponding to Example 3. The covariances of $OU_{\boldsymbol{\kappa}}$ are indicated with a dotted line.*

The associated ARMA(3,2) model is

$$(1-1.9255B+1.05185B^2-0.1200B^3)x=(0.4831-0.9044B+0.4230B^2)\epsilon.$$

Example 3. The parameter $\boldsymbol{\kappa}=(0.83,0.0041,0.0009)$ used in the simulation of the OU process x treated in the present example is approximately equal to the parameter $\hat{\boldsymbol{\kappa}}$ obtained by ML estimation with $p=3$ for Series A in Section 6.5.1. A graphical presentation of the estimated covariances is given in Figure 3. The associated ARMA(3,2) model is

$$(1-2.4311B+1.8649B^2-0.4339B^3)x=(0.6973-1.3935B+0.6962B^2)\epsilon$$

The description of the performance of the model is complemented by comparing in Figure 4 the simulated values of the process in 400 equally spaced points filling the interval (199,201) with the predicted values for the same interval, based on the OU(3) model and the assumed observed data $x(0),x(2),x(3),\ldots,x(200)$. Also a confidence band limited by the predicted values plus and minus twice their standard deviation (2-st.-dev. confidence band) is included in the graph, in order to describe the precision of the predicted values.

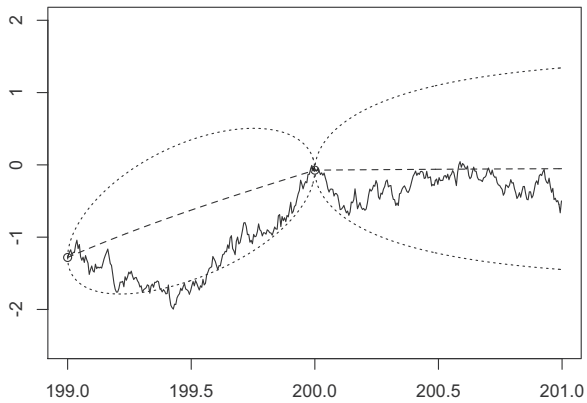


Figure 4: *Estimated interpolation and prediction of* $x(t)$ *for* $199<t<200$ *and* $200<t<201$*, respectively (- - -), 2-st.-dev. confidence bands based on* $(x(i))_{i=0,1,\ldots,200}$ $(\cdots)$*, and a refinement of the simulation of* $x(t)$ *on* $199<t<200$*.*

6.5. Applications to real data

In this section we present experimental results on two real data sets. We fit OU(p) processes for small values of p and also some ARMA processes. In each case we have observed that we can find an adequate value of p for which the empirical covariances are well approximated by the covariances of the adjusted OU(p) model. This is not the case

for the ARMA models adjusted by maximum likelihood, in all examples. We present a detailed comparison of both methodologies for the first example.

The first data set is taken from Box, Jenkins, and Reinsel (1994), and correspond to equally spaced observations of continuous-time processes that might be assumed to be stationary. The second one is a series obtained by choosing one in every 100 terms of a high frequency recording of oxygen saturation in blood of a newborn child. The data were obtained by a team of researchers of Pereira Rossell Children Hospital in Montevideo, Uruguay, integrated by L. Chiapella, A. Criado and C. Scavone. Their permission to analyse the data is gratefully acknowledged by the authors.

6.5.1. Box, Jenkins and Reinsel "Series A"

The Series A is a record of $n = 197$ chemical process concentration readings, taken every two hours, introduced with that name and analysed in (Box, Jenkins, and Reinsel, 1994, Ch. 4)[1]. Box et al. suggest an ARMA(1,1) as a model for this data, and subsets of AR(7) are proposed in (Cleveland, 1971) and (McLeod and Zhang, 2006). Figure 5 shows that these models fit fairly well the autocovariances for small lags, but fail to capture the structure of autocorrelations for large lags present in the series. On the other hand, the approximations obtained with the OU(p) processes, for $p = 3,4$ reflect both the short and long dependences, as shown in Figure 6.

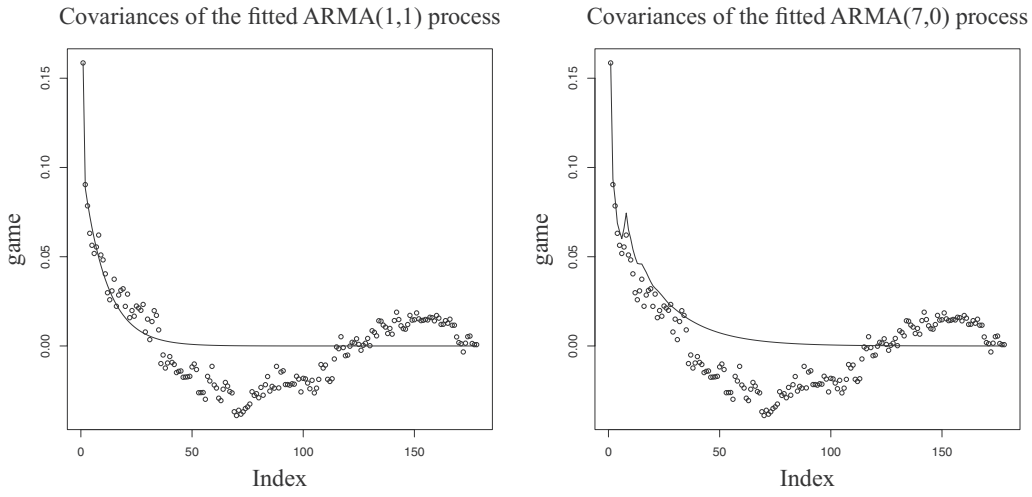


Figure 5: *Empirical covariances (∘) and covariances of the ML (—) fitted models ARMA(1,1) and AR(7) for Series A.*

1. see also http://rgm2.lab.nig.ac.jp/RGM2/tfunc.php?rd_id=FitAR:SeriesA

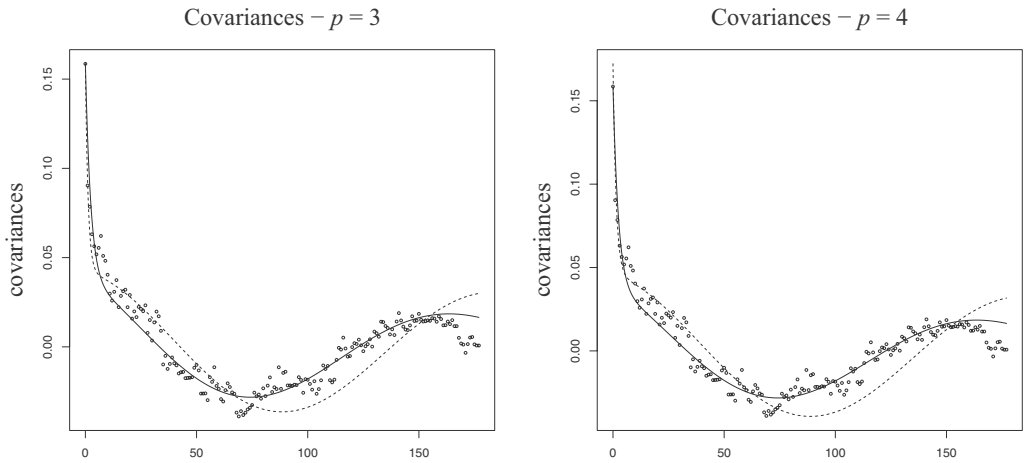


Figure 6: *Empirical covariances (∘) and covariances of the MC (—) and ML (- - -) fitted OU(p) models, for $p = 3,4$ corresponding to Series A.*

It is interesting to consider jointly the ARMA(3,2) model (31) fitted to the original data by maximum likelihood (computed also with the `R` function `arima`) and the ARMA(3,2) model (32) obtained by the procedure described in Section 5, corresponding to the OU(3) process also fitted to the data by maximum likelihood. The estimated parameters of this OU process are

$$\hat{\boldsymbol{\kappa}} = (0.8293, 0.0018 + 0.0330i, 0.0018 - 0.0330i) \text{ and } \hat{c} = 0.4401$$

and the ARMA(3,2) processes are respectively

$$(1-0.7945B-0.3145B^2+0.1553B^3)x=0.3101(1-0.4269B-0.2959B^2)\epsilon \qquad (31)$$

and

$$(1-2.4316B+1.8670B^2-0.4348B^3)x=0.4401(1-1.9675B+0.9685B^2)\epsilon. \qquad (32)$$

The autocorrelations of both ARMA models, shown in Figure 7, together with the empirical correlations of the series were computed by means of the `R` function `ARMAacf`, although the ones corresponding to (32) could have been obtained as the restrictions to integer lags of the covariance function for continuous-time described in Section 3.2. It is worth to notice that the autocorrelations of (31) do not approach the empirical correlations, indicated by circles, as much as the correlations of (32). The logarithms of the likelihoods of (31) and (32) are $\ell' = -49.23$, and $\ell'' = -50.95$, respectively. But since the number of parameters of the second model (which is four) is smaller than the number of parameters of the complete family of ARMA(3,2) processes (six), the Akaike

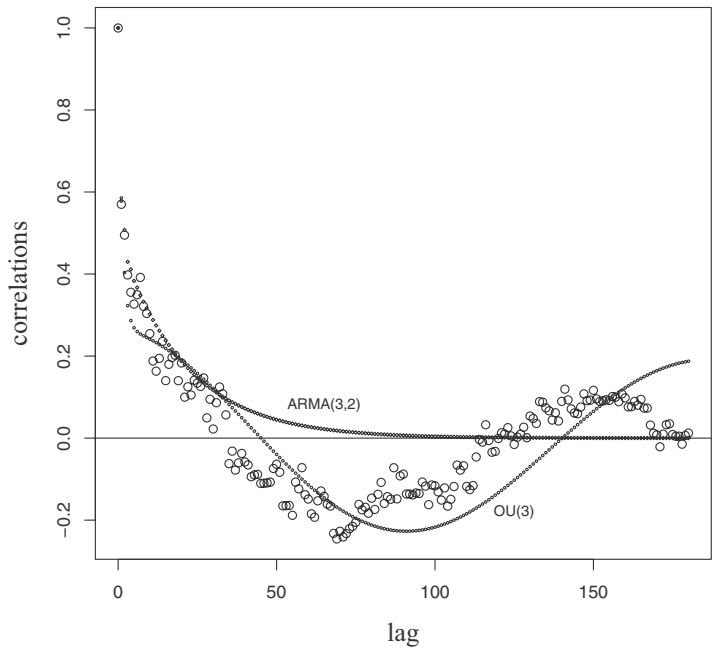


Figure 7: *Empirical correlations (○) of Series A, and autocorrelations of models (31) and (32) fitted by maximum likelihood from the family of all ARMA(3,2) and the restricted family of ARMA(3,2) derived from OU(3).*

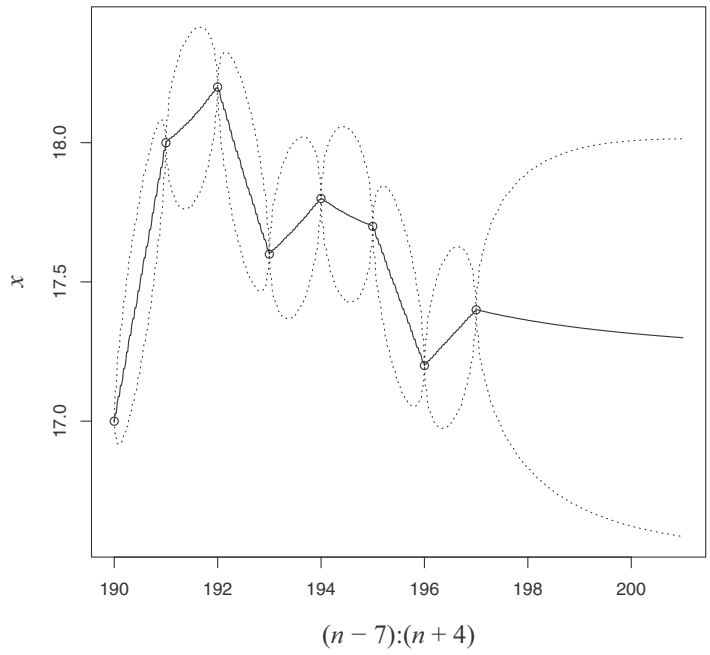


Figure 8: *Confidence bands for interpolated and extrapolated values of Series A for continuous domain.*

information criterion (AIC) of the parsimonious OU model is $8-2\ell'' = 109.90$, slightly better than the AIC of the unrestricted ARMA model, equal to $12-2\ell' = 110.46$.

Finally we show in Figure 8 the predicted values of the continuous parameter process $x(t)$, for t between $n-7$ and $n+4$ (190-201), obtained as the best linear predictions based on the last 90 observed values, and on the correlations given by the fitted OU(3) model. The upper and lower lines are two standard deviation confidence limits for each value of the process.

6.5.2. Oxygen saturation in blood

The oxygen saturation in blood of a newborn child has been monitored during 17 hours, and measures taken every two seconds. We assume that a series $x_0, x_1, \ldots, x_{304}$ of measures taken at intervals of 200 seconds is observed, and fit OU processes of orders $p = 2,3,4$ to that series.

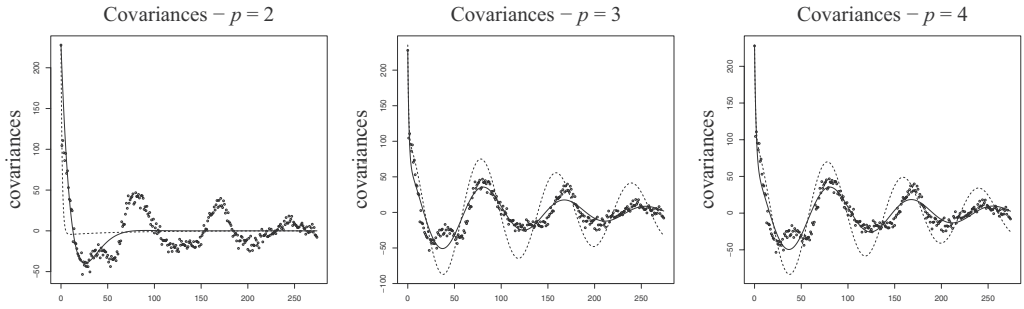


Figure 9: *Empirical covariances (∘) and covariances of the MC (—) and ML (- - -) fitted OU(p) models for $p = 2,3,4$ corresponding to the series of oxygen saturation in blood.*

Again the empirical covariances of the series and the covariances of the fitted OU(p) models for $p = 2,3,4$ are plotted (see Figure 9) and the estimated interpolation and extrapolation are shown in Figure 10. In the present case, the actual values of the series for integer multiples of 1/100 of the unit measure of 200 seconds are known, and plotted in the same figure.

6.6. Estimating the shape of the Lévy noise

There are various methods proposed in the literature to estimate the parameters of Lévy driven Ornstein–Uhlenbeck processes; in particular, the Lévy-Khinchin triplet comprised of two real numbers and a measure. For example, Valdivieso, Schoutens, and Tuerlinckx (2009) propose a maximum likelihood estimation methodology based on the inversion of the characteristic function of the Lévy process and the use of the discrete fast Fourier transform. Jongbloed, van der Meulen, and van der Vaart (2005) propose a nonparametric estimation based on a preliminary estimator of the characteristic function. Both methods require a large amount of information and intensive computation.

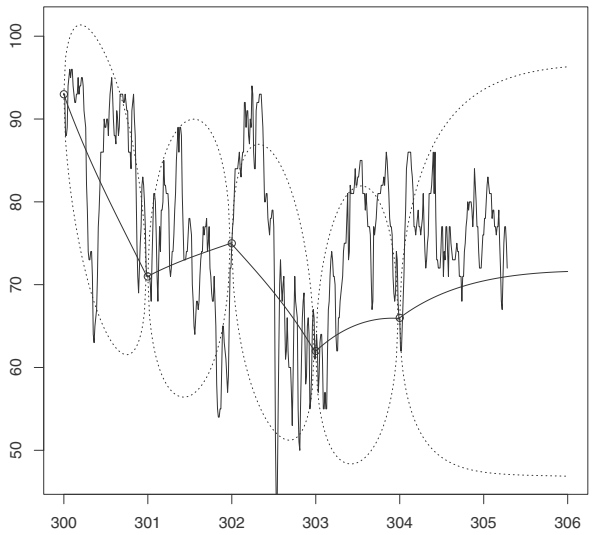


Figure 10: *Partial graph showing the five last values of the series of O_2 saturation in blood at integer multiples of the 200 seconds unit of time (∘), interpolated and extrapolated predictions (—), 2-st.-dev. confidence bands (⋯), and actual values of the series.*

We propose a naive method of estimating the parameters of the Lévy driven Ornstein–Uhlenbeck process that works in general situations when the maximum likelihood function is not known or difficult to approximate. These estimators are easy to compute, but also require a large amount of data to attain high accuracy.

Our method of estimation resembles the methods described in (Yu, 2004) consisting on matching the characteristic function derived from the model and the empirical characteristic function derived from the data.

Given a Lévy process $\Lambda(t)$, the characteristic function of $\Lambda(t)$ is $\mathrm{E}\,\mathrm{e}^{iu\Lambda(t)} = (\mathrm{E}\,\mathrm{e}^{iu\Lambda(1)})^t$, and is usually written as $\mathrm{E}\,\mathrm{e}^{iu\Lambda(1)} = \mathrm{e}^{\psi_\Lambda(iu)}$. The function $\psi_\Lambda(iu) = \log \mathrm{E}\,\mathrm{e}^{iu\Lambda(1)}$ is called *characteristic exponent* and has the form

$$\psi_\Lambda(iu) = aiu - \frac{\sigma^2}{2}u^2 + \int_{|x|<1} (\mathrm{e}^{iux} - 1 - iux)d\nu(x) + \int_{|x|\geq 1} (\mathrm{e}^{iux} - 1)d\nu(x)$$

where $\nu(\{0\}) = 0$, $\int_{|x|<1} x^2 d\nu(x) < \infty$, $\int_{|x|\geq 1} d\nu(x) < \infty$. The Lévy-Khinchin triplet is (σ^2, a, ν).

Assume that the admissible exponents belong to a parametric class $\Psi = \{\psi_\theta : \theta \in \Theta\}$ where $\Theta \subset \mathbb{R}^d$, and obtain the value of θ for which a chosen quadratic distance between the exponential of $\psi_\theta(iu)$ and the empirical characteristic function of the residuals is minimum.

In order to ease notation, let us consider the case of an OU(p) model with parameter $\boldsymbol{\kappa}$ of pairwise different components; either $\boldsymbol{\kappa}$ is known or it is estimated by maximum likelihood or matching correlation methods. The innovation in each component ξ_j is

$$\eta_j(t) = \int_{t-1}^{t} \mathrm{e}^{-\kappa_j(t-s)} d\Lambda(s),$$

so that the innovation of x_κ is

$$\eta(t) = \int_{t-1}^{t} g(t-s) d\Lambda(s) \quad \text{where} \quad g(t) = \sum_{j=1}^{p} K_j \mathrm{e}^{-\kappa_j t}.$$

Hence, if we denote $\eta := \eta(1)$, we have

$$\eta \sim \int_0^1 g(1-s) d\Lambda(s) \sim \int_0^1 g(s) d\Lambda(s)$$

and its characteristic exponent is therefore

$$\psi_\eta(iu) = \log \mathrm{E}\mathrm{e}^{iu\eta} = \log \mathrm{E}\mathrm{e}^{iu\int_0^1 g(s)d\Lambda(s)} = \int_0^1 \psi_\Lambda(iug(s))ds$$

Example 4. Consider the estimation of a noise sum of a Poisson process plus a Gaussian term. Let us assume that the noise is given by

$$\Lambda(t) = \sigma W(t) + a(N(t) - \lambda t)$$

where W is a standard Wiener process and N is a Poisson process with intensity λ. The family of possible noises depends on the three parameters (σ, λ, a). In this case, the characteristic exponent has a simple form:

$$\psi_{\Lambda(1)}(iu) = -\frac{\sigma^2 u^2}{2} + \lambda(\mathrm{e}^{iua} - iua - 1),$$

hence

$$\psi_\eta(iu) = \int_0^1 \left(-\frac{\sigma^2 u^2 g^2(s)}{2} + \lambda(\mathrm{e}^{iug(s)a} - iug(s)a - 1) \right) ds$$

Defining $g_h = \int_0^1 g^h(s)ds$, we have

$$\psi_\eta(iu) = -\frac{\sigma^2 u^2 g_2}{2} + \lambda \left(-\frac{u^2 g_2 a^2}{2} - i\frac{u^3 g_3 a^3}{6} + \frac{u^4 g_4 a^4}{24} + \dots \right)$$

Then we propose to estimate the parameters by equating the coefficients of u^2, u^3, u^4 in $\psi_\eta(iu)$ with the corresponding ones in the logarithm of the empirical characteristic function of the residuals.

Assuming that the mean of the residuals $r_1, r_2, \ldots, r_n$ is zero, their empirical characteristic function is

$$\frac{1}{n}\sum_{h=1}^{n} \mathrm{e}^{iur_h} = 1 - \frac{1}{2}u^2 R_2 - \frac{1}{6}iu^3 R_3 + \frac{1}{24}u^4 R_4 + \ldots$$

where $R_m = \frac{1}{n}\sum_{h=1}^{n} r_h^m$. Then the logarithm has the expansion

$$\log \frac{1}{n}\sum_{h=1}^{n} \mathrm{e}^{iur_h} = -\frac{1}{2}u^2 R_2 - \frac{1}{6}iu^3 R_3 + \frac{1}{24}u^4 R_4 - \frac{1}{8}u^4 R_2^2 + \ldots$$

Consequently, the estimation equations are

$$\begin{aligned} (\sigma^2 + \lambda a^2) g_2 &= R_2, \\ \lambda a^3 g_3 &= R_3, \\ \lambda a^4 g_4 &= R_4 - 3R_2^2 \end{aligned}$$

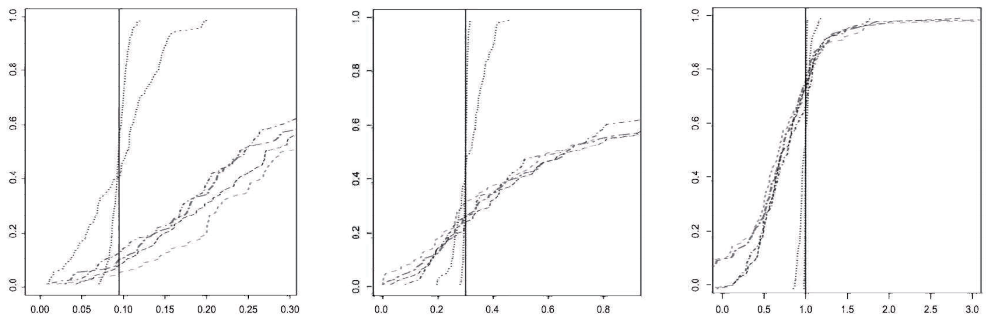
from which the estimators follow:

$$\tilde{a} = \frac{R_4 - 3R_2^2}{R_3}\frac{g_3}{g_4}, \qquad \tilde{\lambda} = \frac{R_3^4}{(R_4 - 3R_2^2)^3}\frac{g_4^3}{g_3^4},$$

$$\tilde{\sigma}^2 = \frac{R^2}{g_2} - \frac{R_3^2}{(R_4 - 3R_2^2)}\frac{g_4}{g_3^2}.$$

Figure 11 shows the empirical c.d.f. of 90 estimators of the parameters obtained from simulated series of 200 terms. The residuals were obtained by applying a Kalman filter to the space state formulation, starting from the actual value of $\boldsymbol{\kappa}$ used at the simulation (-·-), that in practical situations is unknown, and from matching correlations estimation (– –) and by maximum likelihood estimation (–·–).

The estimators are not sharp at all, but the ones obtained by the same procedure applied directly on the unfiltered noise Λ (– –) are equally rough. Larger series (of size 10000 and 1000000) produce sharper estimates, also shown in the figures by dotted lines.



c.d.f. of 90 estimators of σ *c.d.f. of 90 estimators of λ* *c.d.f. of 90 estimators of c*

Figure 11: *Estimation of the parameters of the noise (σ –left panel–, λ –center–, a –right–) from 90 replications of* $\{x_{\boldsymbol{\kappa}}(t) : t = 0,1,\ldots,200\}$, $\boldsymbol{\kappa} = (0.01 \pm 0.1i, 0.2)$, *driven by* $\Lambda(t) = 0.1W(t) + N_{0.3}(t) - 0.3t$. *Normality is rejected in all cases.*

7. Conclusions

We have proposed a family of continuous-time stationary processes, based on p iterations of the linear operator that maps a second order Lévy process onto an Ornstein-Uhlenbeck process. These operators have some nice properties, such as being commutative, and their p-compositions decompose as a linear combination of simple operators of the same kind. We remark that this result, stated in Theorem 1, is independent of the process onto which the operators $\mathcal{OU}_{\boldsymbol{\kappa}}$ act on. We have reduced the present scope of the applications envisaged by applying the operators only to Lévy processes, but other choices deserve consideration, for example, the results of applying the same operators to fractional Brownian motions.

An OU(p) process depends on $p+1$ parameters that can be easily estimated by either maximum likelihood (ML) or matching correlations (MC) procedures. MC estimators provide a fair estimation of the covariances of the data, even if the model is not well specified. When sampled on equally spaced instants, the OU(p) family can be written as a discrete time state-space model; i.e., a VARMA model in a space of dimension p. As a consequence, the families of OU(p) models are a parsimonious subfamily of the ARMA$(p, p-1)$ processes in the Gaussian case. Furthermore, the coefficients of the ARMA can be deduced from those of the corresponding OU(p). We have shown examples for which the ML-estimated OU model is able to capture features of the empirical autocorrelations at large lags that the ML-estimated ARMA model does not (see for instance Figure 7). This leads to recommend the inclusion of OU models as candidates to represent stationary series, either in discrete time or continuous-time.

References

Barndorff-Nielsen, O.E. (2001). Superposition of Ornstein-Uhlenbeck type processes. *Theory of Probability and Its Applications*, 45, 175–194.

Bergstrom, A.R. (1984). Continuous time stochastic models and issues of aggregation over time. *Handbook of Econometrics*, II, 1145–1212.

Bergstrom, A.R. (1996). Survey of continuous-time econometrics. In *Dynamic Disequilibrium Modeling: Theory and Applications: Proceedings of the Ninth International Symposium in Economic Theory and Econometrics*, volume 9, page 1. Cambridge University Press.

Box, G.E.P. Jenkins, G.M. and Reinsel, G.C. (1994). *Time Series Analysis, Forecasting and Control*. Prentice Hall.

Brockwell, P.J. (2004). Representations of continuous-time ARMA processes. *Journal of Applied Probability*, 41, 375–382.

Brockwell, P.J. (2009). Lévy–driven continuous–time ARMA processes. In *Handbook of Financial Time Series*, pages 457–480. Springer.

Cleveland, W.S. (1971). The inverse autocorrelations of a time series and their applications. *Technometrics*, 14, 277–298.

Doob, J.L. (1944). The elementary Gaussian processes. *Annals of Mathematical Statistics*, 15, 229–282.

Durbin, J. (1961). Efficient fitting of linear models for continuous stationary time-series from discrete data. *Bulletin of the International Statistical Institute*, 38, 273–282.

Durbin, J. and Koopman, S.J. (2001). *Time Series Analysis by State Space Methods*. Oxford University Press.

Eliazar, I. and Klafter, J. (2009). From Ornstein-Uhlenbeck dynamics to long-memory processes and Fractional Brownian motion. *Physical Review E*, 79, 021115.

Granger, C.W.J. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, 14, 227–238.

Granger, C.W.J. and Morris, M.J. (1976). Time series modelling and interpretation. *Journal of the Royal Statistical Society. Series A*, 139, 246–257.

Jongbloed, G., van der Meulen, F.H. and van der Vaart, A.W. (2005). Nonparametric inference for Lévy-driven Ornstein-Uhlenbeck processes. *Bernoulli*, 11, 759–791.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer Science & Business Media.

McLeod, A.I. and Zhang, Y. (2006). Partial autocorrelation parameterization for subset autoregression. *Journal of Time Series Analysis*, 27, 599–612.

Nieto, B., Orbe, S. and Zarraga, A. (2014). Time-Varying Market Beta: Does the estimation methodology matter? *SORT*, 31, 13–42.

R Core Team. (2015). *R: A Language and Environment for Statistical Computing*. Technical report, R Foundation for Statistical Computing, Vienna, Austria.

Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distribution*, volume 68 of *Cambridge Studies in Advance Mathematics*. Cambridge University Press.

Thornton, M.A. and Chambers, M.J. (2013). Continuous-time autoregressive moving average processes in discrete time: representation and embeddability. *Journal of Time Series Analysis*, 34, 552–561.

Uhlenbeck, G.E. and Ornstein, L.S. (1930). On the theory of the Brownian motion. *Physical Review*, 36, 823–841.

Valdivieso, L., Schoutens, W. and Tuerlinckx, F. (2009). Maximum likelihood estimation in processes of Ornstein-Uhlenbeck type. *Statistical Inference for Stochastic Processes*, 12, 1–19.

Yu, J. (2004). Empirical characteristic function estimation and its applications. *Econometric Reviews*, 23, 93–123.

Appendix A: Proofs of Theorem 1 and its corollaries

Parts (i) and (iii) are obtained by direct computation of the integrals, (ii) follows from (i) by finite induction, as well as (iv) from (iii).

From the continuity of the integrals with respect to the parameter κ, the power $\mathcal{OU}_\kappa^p$ satisfies

$$\mathcal{OU}_\kappa^p = \lim_{\delta\downarrow 0}\prod_{j=1}^{p}\mathcal{OU}_{\kappa+j\delta} = \lim_{\delta\downarrow 0}\sum_{j=1}^{p}K'_j(\delta,\kappa,p)\mathcal{OU}_{\kappa+j\delta} \tag{33}$$

with

$$K'_j(\delta,\kappa,p) = \frac{1}{\prod_{1\le l\le p, l\neq j}\left(1-\frac{\kappa+l\delta}{\kappa+j\delta}\right)}.$$

On the other hand, by (i),

$$\prod_{h=1}^{q}\mathcal{OU}_{\kappa_h}^{p_h} = \lim_{\boldsymbol{\delta}\downarrow 0}\prod_{h=1}^{q}\prod_{j=1}^{p_h}\mathcal{OU}_{\kappa_h+j\delta_h} = \lim_{\boldsymbol{\delta}\downarrow 0}\sum_{h=1}^{q}\sum_{j=1}^{p_h}K''_{h,j}(\boldsymbol{\delta},\boldsymbol{\kappa})\mathcal{OU}_{\kappa_h+j\delta_h} \tag{34}$$

where $\boldsymbol{\delta} = (\delta_1,\ldots,\delta_q)$,

$$K''_{h,j}(\boldsymbol{\delta},\boldsymbol{\kappa}) = \frac{1}{\prod_{\substack{1\le h'\le q, 1\le j'\le p_h,\\ (h',j')\neq(h,j)}}\left(1-\frac{\kappa_{h'}+j'\delta_{h'}}{\kappa_h+j\delta_h}\right)} = K'''_{h,j}(\boldsymbol{\delta},\boldsymbol{\kappa})K'_j(\delta_h,\kappa_h,p_h),$$

and

$$K'''_{h,j}(\boldsymbol{\delta},\boldsymbol{\kappa}) = \frac{1}{\prod_{\substack{1\le h'\le q,\\ h'\neq h}}\prod_{j'=1}^{p_{h'}}(1-(\kappa_{h'}+j'\delta_{h'})/(\kappa_h+j\delta_h))} \to K_h(\boldsymbol{\kappa}) \text{ as } \boldsymbol{\delta}\downarrow 0$$

For the h-th term in the right-hand side of (34), we compute

$$\begin{aligned}\lim_{\boldsymbol{\delta}\downarrow 0}\sum_{j=1}^{p_h}K''_{h,j}(\boldsymbol{\delta},\boldsymbol{\kappa})\mathcal{OU}_{\kappa_h+j\delta_h} &= \lim_{\boldsymbol{\delta}\downarrow 0}\sum_{j=1}^{p_h}K'''_{h,j}(\boldsymbol{\delta},\boldsymbol{\kappa})K'_j(\delta_h,\kappa_h,p_h)\mathcal{OU}_{\kappa_h+j\delta_h}\\ &= \lim_{\boldsymbol{\delta}\downarrow 0}\sum_{j=1}^{p_h}(K'''_{h,j}(\boldsymbol{\delta},\boldsymbol{\kappa})-K_h(\boldsymbol{\kappa}))K'_j(\delta_h,\kappa_h,p_h)\mathcal{OU}_{\kappa_h+j\delta_h}\\ &+\ K_h(\boldsymbol{\kappa})\lim_{\boldsymbol{\delta}\downarrow 0}\sum_{j=1}^{p_h}K'_j(\delta_h,\kappa_h,p_h)\mathcal{OU}_{\kappa_h+j\delta_h} = K_h(\boldsymbol{\kappa})\mathcal{OU}_{\kappa_h}^{p_h}\end{aligned}$$

by Equation (33) since, in addition, each term in the first sum tends to zero. This ends the verification of (v).

Corollary 1 is an immediate consequence of (iv) and (v), and Corollary 2 follows by applying (i) to compute

$$\begin{aligned}\mathcal{OU}_{\lambda+i\mu}\mathcal{OU}_{\lambda-i\mu} &= \frac{\lambda+i\mu}{2i\mu}\mathcal{OU}_{\lambda+i\mu} - \frac{\lambda-i\mu}{2i\mu}\mathcal{OU}_{\lambda-i\mu} \\ &= \int_{-\infty}^{t} \mathrm{e}^{-\lambda(t-s)} \left[\tfrac{\lambda+i\mu}{2i\mu}(\cos(\mu(t-s)) + i\sin(\mu(t-s))) \right. \\ &\quad \left. -\tfrac{\lambda-i\mu}{2i\mu}(\cos(\mu(t-s)) - i\sin(\mu(t-s)))\right] d\Lambda(s) \\ &= \int_{-\infty}^{t} \mathrm{e}^{-\lambda(t-s)}(\cos(\mu(t-s)) + \tfrac{\lambda}{\mu}\sin(\mu(t-s)))d\Lambda(s). \qquad \blacksquare\end{aligned}$$

Appendix B: Derivation of a state-space model

The form of equations (22) for a state-space representation of the OU(p) process in the general case can be derived by considering three special cases:

1. *When the components of $\boldsymbol{\kappa}$ are all different*. This case is treated in Section 4.
2. *When the components of $\boldsymbol{\kappa}$ are all equal*. Let κ denote the common value of the components of $\boldsymbol{\kappa}$. The state of the system is described by the vector

$$\boldsymbol{\xi}_{\kappa,p} = (\xi_\kappa^{(0)}, \xi_\kappa^{(1)}, \ldots, \xi_\kappa^{(p-1)})^\intercal,$$

with components $\xi_\kappa^{(h)}(t) = \displaystyle\int_{-\infty}^{t} \mathrm{e}^{-\kappa(t-s)} \frac{(-\kappa(t-s))^h}{h!} d\Lambda(s)$.
Each of these terms can be written as the sum

$$\xi_\kappa^{(h)}(t) = \mathrm{e}^{-\kappa} \int_{-\infty}^{t-1} \mathrm{e}^{-\kappa(t-1-s)} \frac{(-\kappa(t-1-s+1))^h}{h!} d\Lambda(s) + \eta_{\kappa,h}(t) \tag{35}$$

where $\eta_{\kappa,h}(t) = \displaystyle\int_{t-1}^{t} \mathrm{e}^{-\kappa(t-s)} \frac{(-\kappa(t-s))^h}{h!} d\Lambda(s)$.
The first term in the right-hand side of (35) is equal to

$$\begin{aligned}&\mathrm{e}^{-\kappa} \sum_{j=0}^{h} \frac{(-\kappa)^{h-j}}{(h-j)!} \int_{-\infty}^{t-1} \mathrm{e}^{-\kappa(t-1-s)} \frac{(-\kappa(t-1-s))^j}{j!} d\Lambda(s) \\ &= \mathrm{e}^{-\kappa} \sum_{j=0}^{h} \frac{(-\kappa)^{h-j}}{(h-j)!} \xi_\kappa^{(j)}(t-1)\end{aligned}$$

and therefore, by introducing the matrix

$$\boldsymbol{A}_{\kappa,p} = \mathrm{e}^{-\kappa} \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ \frac{(-\kappa)}{1!} & 1 & 0 & \dots & 0 & 0 \\ \frac{(-\kappa)^2}{2!} & \frac{(-\kappa)}{1!} & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{(-\kappa)^{p-2}}{(p-2)!} & \frac{(-\kappa)^{p-3}}{(p-3)!} & \frac{(-\kappa)^{p-4}}{(p-4)!} & \dots & 1 & 0 \\ \frac{(-\kappa)^{p-1}}{(p-1)!} & \frac{(-\kappa)^{p-2}}{(p-2)!} & \frac{(-\kappa)^{p-3}}{(p-3)!} & \dots & \frac{(-\kappa)}{1!} & 1 \end{pmatrix}$$

we may write

$$\boldsymbol{\xi}_{\kappa,p}(t) = \boldsymbol{A}_{\kappa,p}\boldsymbol{\xi}_{\kappa,p}(t-1) + \boldsymbol{\eta}_{\kappa,p}$$

where $\boldsymbol{\eta}_{\kappa,p}(t) = (\eta_{\kappa,0}(t), \eta_{\kappa,1}(t), \dots, \eta_{\kappa,p-1}(t))^\top$ is a vector of centered innovations (independent of the σ-algebra generated by $\{\Lambda(s) : s \le t-1\}$) with covariance matrix $\boldsymbol{B}_{\kappa,\kappa,p}$ obtained with $\kappa_1 = \kappa_2$ and $p_1 = p_2$ from the general expression of the $p_1 \times p_2$ matrix $B_{\kappa_1,\kappa_2,p_1,p_2} = ((b_{\kappa_1,\kappa_2,h_1,h_2}))_{1\le h_1\le p_1, 1\le h_2\le p_2}$, where

$$\begin{aligned} b_{\kappa_1,\kappa_2,h_1,h_2} &= \mathrm{E}\,\eta_{\kappa_1,h_1}(t)\bar{\eta}_{\kappa_2,h_2}(t) \\ &= v^2 \int_{t-1}^{t} \mathrm{e}^{-(\kappa_1+\bar{\kappa}_2)(t-s)}(-\kappa_1)^{h_1}(-\bar{\kappa}_2)^{h_2}(t-s)^{h_1+h_2}ds \\ &= v^2 \int_0^1 \mathrm{e}^{-(\kappa_1+\bar{\kappa}_2)y}(-\kappa)^{h_1}(-\bar{\kappa})^{h_2}y^{h_1+h_2}dy. \end{aligned}$$

The equation $x(t) = \boldsymbol{K}_p^\top \boldsymbol{\xi}_{\boldsymbol{\kappa},p}(t)$, with $\boldsymbol{K}_p^\top = (\binom{p-1}{0}, \binom{p-1}{1}, \dots, \binom{p-1}{p-1})$ completes the description of the system state dynamics.

3. *The vector κ has components $\kappa_1 = \lambda + \mu i$ and $\kappa_2 = \lambda - \mu i$, $\mu \neq 0$, each repeated p_1 times*. A description involving imaginary processes is immediate from the previous case. The equations

$$\begin{pmatrix} \boldsymbol{\xi}_{\kappa_1,p_1}(t) \\ \boldsymbol{\xi}_{\kappa_2,p_1}(t) \end{pmatrix} = \begin{pmatrix} A_{\kappa_1,p_1} & 0 \\ 0 & A_{\kappa_2,p_1} \end{pmatrix} \begin{pmatrix} \boldsymbol{\xi}_{\kappa_1,p_1}(t-1) \\ \boldsymbol{\xi}_{\kappa_2,p_1}(t-1) \end{pmatrix} + \begin{pmatrix} \boldsymbol{\eta}_{\kappa_1,p_1} \\ \boldsymbol{\eta}_{\kappa_2,p_1} \end{pmatrix}$$

$$x(t) = (\boldsymbol{K}_{p_1}^\top, \boldsymbol{K}_{p_1}^\top) \begin{pmatrix} \boldsymbol{\xi}_{\kappa_1,p_1}(t) \\ \boldsymbol{\xi}_{\kappa_2,p_1}(t) \end{pmatrix}$$

hold, and $\mathrm{Var}\begin{pmatrix} \boldsymbol{\eta}_{\kappa_1,p_1} \\ \boldsymbol{\eta}_{\kappa_2,p_1} \end{pmatrix} = \begin{pmatrix} \boldsymbol{B}_{\kappa_1,\kappa_1,p_1,p_1} & \boldsymbol{B}_{\kappa_1,\kappa_2,p_1,p_1} \\ \boldsymbol{B}_{\kappa_2,\kappa_1,p_1,p_1} & \boldsymbol{B}_{\kappa_1,\kappa_1,p_1,p_1} \end{pmatrix}$.

A description in terms of real processes is obtained by multiplying the first equation by the matrix

$$\boldsymbol{C}_{p_1} = \begin{pmatrix} I_{p_1} & I_{p_1} \\ -iI_{p_1} & iI_{p_1} \end{pmatrix}$$

(I_p denotes the $p \times p$ identity matrix), because the vectorial process $\boldsymbol{C}_{p_1} \begin{pmatrix} \boldsymbol{\xi}_{\kappa_1,p_1}(t) \\ \boldsymbol{\xi}_{\kappa_2,p_1}(t) \end{pmatrix}$ has real components. The new equations are

$$\boldsymbol{C}_{p_1} \begin{pmatrix} \boldsymbol{\xi}_{\kappa_1,p_1}(t) \\ \boldsymbol{\xi}_{\kappa_2,p_1}(t) \end{pmatrix} = \boldsymbol{C}_{p_1} \begin{pmatrix} \boldsymbol{\eta}_{\kappa_1,p_1} \\ \boldsymbol{\eta}_{\kappa_2,p_1} \end{pmatrix} + \left(\boldsymbol{C}_{p_1} \begin{pmatrix} \boldsymbol{A}_{\kappa_1,p_1} & 0 \\ 0 & \boldsymbol{A}_{\kappa_2,p_1} \end{pmatrix} \boldsymbol{C}_{p_1}^{-1} \right) \times \left(\boldsymbol{C}_{p_1} \begin{pmatrix} \boldsymbol{\xi}_{\kappa_1,p_1}(t-1) \\ \boldsymbol{\xi}_{\kappa_2,p_1}(t-1) \end{pmatrix} \right)$$

and

$$\boldsymbol{x}(t) = \left((\boldsymbol{K}_{p_1}^{\top}, \boldsymbol{K}_{p_1}^{\top}) C_{p_1}^{-1} \right) \times \left(\boldsymbol{C}_{p_1} \begin{pmatrix} \boldsymbol{\xi}_{\kappa_1,p_1}(t) \\ \boldsymbol{\xi}_{\kappa_2,p_1}(t) \end{pmatrix} \right)$$

General case, real processes

Let us assume that $\kappa_1, \ldots, \kappa_q$ are distinct components of $\boldsymbol{\kappa}$, each repeated $p_1, \ldots, p_q$ times. We assume in addition that the imaginary components are $\kappa_1, \kappa_2 = \bar{\kappa}_1, \ldots, \kappa_{2c-1}$, $\kappa_{2c} = \bar{\kappa}_{2c-1}$ and the remaining $\kappa_{2c+1}, \ldots, \kappa_q$ are real. With this notation, $p_{2h-1} = p_{2h}$ for $h = 1, 2, \ldots, c$. We make intensive use of the notations introduced in previous cases to write

$$\boldsymbol{\xi}(t) = \boldsymbol{A}\boldsymbol{\xi}(t-1) + \boldsymbol{\eta}(t), \tag{36}$$

$$\boldsymbol{x}(t) = K^{\top}\boldsymbol{\xi}(t)$$

with

$$\boldsymbol{\xi}(t) = \begin{pmatrix} \boldsymbol{\xi}_{\kappa_1,p_2}(t) \\ \boldsymbol{\xi}_{\kappa_2,p_2}(t) \\ \boldsymbol{\xi}_{\kappa_3,p_4}(t) \\ \boldsymbol{\xi}_{\kappa_4,p_4}(t) \\ \cdots \\ \boldsymbol{\xi}_{\kappa_{2c-1},p_{2c}}(t) \\ \boldsymbol{\xi}_{\kappa_{2c},p_{2c}}(t) \\ \boldsymbol{\xi}_{\kappa_{2c+1},p_{2c+1}}(t) \\ \boldsymbol{\xi}_{\kappa_{2c+2},p_{2c+2}}(t) \\ \cdots \\ \boldsymbol{\xi}_{\kappa_q,p_q}(t) \end{pmatrix}, \boldsymbol{\eta}(t) = \begin{pmatrix} \boldsymbol{\eta}_{\kappa_1,p_2}(t) \\ \boldsymbol{\eta}_{\kappa_2,p_2}(t) \\ \boldsymbol{\eta}_{\kappa_3,p_4}(t) \\ \boldsymbol{\eta}_{\kappa_4,p_4}(t) \\ \cdots \\ \boldsymbol{\eta}_{\kappa_{2c-1},p_{2c}}(t) \\ \boldsymbol{\eta}_{\kappa_{2c},p_{2c}}(t) \\ \boldsymbol{\eta}_{\kappa_{2c+1},p_{2c+1}}(t) \\ \boldsymbol{\eta}_{\kappa_{2c+2},p_{2c+2}}(t) \\ \cdots \\ \boldsymbol{\eta}_{\kappa_q,p_q}(t) \end{pmatrix},$$

$$\boldsymbol{A} = \begin{pmatrix} A_{\kappa_1,p_2} & 0 & \dots & 0 \\ 0 & A_{\kappa_2,p_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & A_{\kappa_q,p_q} \end{pmatrix}$$

and

$$\boldsymbol{K}^\top = (K^\top_{\kappa_1,p_1}, K^\top_{\kappa_2,p_2}, \dots, K^\top_{\kappa_q,p_q}).$$

The real version, when the process $\boldsymbol{\xi}$ has imaginary components is obtained by multiplying (36) by the matrix

$$\boldsymbol{C} = \begin{pmatrix} C_{p_2} & 0 & \dots & 0 & 0 \\ 0 & C_{p_4} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & C_{p_{2c}} & 0 \\ 0 & 0 & \dots & 0 & I_{p_{2c+1}+\dots+p_q} \end{pmatrix} \tag{37}$$

thus obtaining

$$\boldsymbol{C\xi}(t) = (\boldsymbol{CAC}^{-1}) \times (\boldsymbol{C\xi}(t-1)) + \boldsymbol{C\eta}(t), \tag{38}$$

$$\boldsymbol{x}(t) = (\boldsymbol{K}^\top \boldsymbol{C}^{-1}) \times (\boldsymbol{C\xi}(t)). \tag{39}$$

Appendix C: Identification of the ARMA

In order to find the coefficients of the ARMA with the same autocovariances as $x(t)$ in the general case, we need the following technical results.

Lemma 1 *For each positive integer p,* $\sum_{j=1}^{p} j^{p-1} \prod_{l \neq j} \frac{1}{j-l} = 1$ *and for* $h = 0, 1, \dots, p-2$, $\sum_{j=1}^{p} j^h \prod_{l \neq j} \frac{1}{j-l} = 0$.

Proof: The polynomial $G(z) = \sum_{j=1}^{p} \left(\frac{1}{j}\right)^{p-1-h} \prod_{l \neq j} \frac{1-lz}{1-l/j}$ has degree $p-1$ and coincides for p different values of the variable, namely $z = 1/j$, $j = 1, 2, \dots, p$, with the polynomial z^{p-1-h}, also of degree not greater than $p-1$ for $h = 0, 1, \dots, p-1$. Therefore, both polynomials are identical, and hence $G(0) = 0$ for $h < p-1$ and $G(0) = 1$ for $h = p-1$. ∎

Lemma 2 *The power series* $g(z,n) = \sum_{h=0}^{\infty} z^h h^n$, $|z| < 1, n = 0,1,2,\ldots$ *has the sum*

$$\sum_{h=0}^{n} \alpha_{n,h}(1-z)^{-h-1}$$

with coefficients determined by $\alpha_{0,0} = 1$ *and the recurrence relations*

$$\alpha_{n+1,h} = h\alpha_{n,h-1} - (h+1)\alpha_{n,h}, h = 0,1,\ldots,n+1, n = 0,1,2,\ldots, \alpha_{n,n+1} = 0.$$

In particular, $\alpha_{n,0} = (-1)^n$. ∎

As an intermediate step from the case described in Section 5.1 and building to the general case, let us approach the OU(p) process x with parameter equal to the p-vector with equal components $\boldsymbol{\kappa} = (\kappa,\kappa,\ldots,\kappa)^\intercal$ as the limit of $x_\delta = \mathcal{OU}_{\boldsymbol{\kappa}(\delta)}\Lambda$, $\boldsymbol{\kappa}(\delta) = (\kappa(1+\delta),\kappa(1+2\delta),\ldots,\kappa(1+p\delta))^\intercal$ when δ tends to zero. From the results in Section 5.1 we use the representation

$$x_\delta = \sum_{j=1}^{p} K_j \xi_j, \quad K_j = \frac{(1+j\delta)^{p-1}}{\delta^{p-1}} \prod_{l \neq j} \frac{1}{j-l} \tag{40}$$

in terms of the vector

$$\boldsymbol{\xi} = (\xi_1,\xi_2,\ldots,\xi_p)^\intercal, \ \ \xi_j(t) = \int_{-\infty}^{t} \mathrm{e}^{-\kappa(1+j\delta)(t-s)} d\Lambda(s)$$

that satisfies $\boldsymbol{\xi} = \mathrm{diag}(\mathrm{e}^{-\kappa(1+j\delta)})B\boldsymbol{\xi} + \boldsymbol{\eta}$ where B is the backshift operator defined in Section 2 and

$$\eta_j(t) = \int_{t-1}^{t} \mathrm{e}^{-\kappa(1+j\delta)(t-s)} d\Lambda(s)$$

and introduce the power expansions

$$\xi_j(t) = \int_{-\infty}^{t} \mathrm{e}^{-\kappa(t-s)} \sum_{h=0}^{\infty} \frac{(j\delta)^h(-\kappa(t-s))^h}{h!} d\Lambda(s) = \sum_{h=0}^{\infty} (j\delta)^h \xi_\kappa^{(h)}(t)$$

with $\xi_\kappa^{(h)}(t) = \int_{-\infty}^{t} \mathrm{e}^{-\kappa(t-s)} \frac{(-\kappa(t-s))^h}{h!} d\Lambda(s)$ and the similar expansion for the innovations

$$\eta_j(t) = \sum_{h=0}^{\infty} (j\delta)^h \eta_\kappa^{(h)}(t) \text{ with } \eta_\kappa^{(h)}(t) = \int_{t-1}^{t} \mathrm{e}^{-\kappa(t-s)} \frac{(-\kappa(t-s))^h}{h!} d\Lambda(s). \tag{41}$$

We write now the ARMA model

$$\prod_{j=1}^{p}(1-\mathrm{e}^{-\kappa(1+j\delta)}B)x_\delta = \sum_{j=1}^{p}\prod_{l\neq j}(1-\mathrm{e}^{-\kappa(1+l\delta)}B)K_j\eta_j$$

and notice that the limit when $\delta \to 0$ of the left-hand side is $(1-\mathrm{e}^{-\kappa}B)^p x$.

In order to take limits at the right-hand side, we replace K_j by its expression in (40), expand $\prod_{l\neq j}(1-\mathrm{e}^{-\kappa(1+l\delta)}B)$ as the product of the series

$$\prod_{l=1}^{p}(1-\mathrm{e}^{-\kappa(1+l\delta)}B) = \sum_{\nu=0}^{\infty}a_\nu\delta^\nu \tag{42}$$

independent of j and

$$(1-\mathrm{e}^{-\kappa(1+j\delta)}B)^{-1} = \sum_{h=0}^{\infty}(\mathrm{e}^{-\kappa(1+j\delta)}B)^h = \sum_{\mu=0}^{\infty}b_\mu(j\delta)^\mu \tag{43}$$

with coefficients independent of j and substitute the expansion (41) for η_j thus obtaining the series

$$\sum_{j=1}^{p}\left(\sum_{\nu=0}^{\infty}a_\nu\delta^\nu \times \sum_{\mu=0}^{\infty}b_\mu(j\delta)^\mu \times (1+j\delta)^{p-1}\prod_{l\neq j}\frac{1}{j-l} \times \sum_{h=0}^{\infty}(j\delta)^h\eta_\kappa^{(h)}\right)$$

divided by δ^{p-1}. After ordering this series by increasing powers of δ, it may be noticed that the terms in δ raised to a power smaller than $p-1$ vanish, because their coefficient include a factor $\sum_{j=1}^{p} j^h \prod_{l\neq j}\frac{1}{j-l}$ with $h \in \{0,1,\ldots,p-2\}$ that is equal to zero as established in Lemma 1 below. Therefore, the limit when $\delta \to 0$ of the series divided by δ^{p-1} is the coefficient of δ^{p-1} in the series. Unless the term a_0 of the first factor is taken, the power of j appearing in the coefficient of δ^{p-1} will be smaller than $p-1$ and again Lemma 1 leads to conclude that the coefficient vanishes. Therefore, since the same lemma establishes that $\sum_{j=1}^{p} j^{p-1}\prod_{l\neq j}\frac{1}{j-l} = 1$, the required limit is the linear combination of moving averages

$$a_0 \sum_{\mu+i+h=p-1}\binom{p-1}{i}b_\mu\eta_\kappa^{(h)} \tag{44}$$

where it remains to make explicit the dependence with respect to the backshift operator B.

From (42) it follows immediately that $a_0 = (1-\mathrm{e}^{-\kappa}B)^p$, while from (43) we get $b_\mu j^\mu \mu! = \left[\frac{\partial^\mu}{\partial\delta^\mu}\sum_{h=0}^{\infty}\mathrm{e}^{-\kappa h}B^h\mathrm{e}^{-hj\delta}\right]_{\delta=0} = (-j)^\mu\sum_{h=0}^{\infty}(\mathrm{e}^{-\kappa}B)^h h^\mu$ and hence

$$b_\mu = \frac{(-1)^\mu}{\mu!}\sum_{\nu=0}^{\infty}(\mathrm{e}^{-\kappa}B)^\nu\nu^\mu.$$

Now we apply Lemma 2 (stated at the end of this section) such that, with the coefficients $\alpha_{\mu,\nu}$ there defined, leads us to write

$$\sum_{\nu=0}^{\infty}(\mathrm{e}^{-\kappa}B)^\nu\nu^\mu = \sum_{\nu=0}^{\mu}\alpha_{\mu,\nu}(1-\mathrm{e}^{-\kappa}B)^{-\nu-1}$$

and therefore (44) is equal to the moving average of order at most $p-1$

$$\sum_{\mu+i+h=p-1}\binom{p-1}{i}\frac{(-1)^\mu}{\mu!}\sum_{\nu=0}^{\mu}\alpha_{\mu,\nu}(1-\mathrm{e}^{-\kappa}B)^{p-\nu-1}\eta_\kappa^{(h)}. \tag{45}$$

Let us observe finally that the order of the moving average is actually $p-1$. The term in B^{p-1} corresponds to $\nu=0$ and reduces to

$$\sum_{\mu+i+h=p-1}\binom{p-1}{i}\frac{(-1)^\mu}{\mu!}\alpha_{\mu,0}(-1)^{p-1}e^{-(p-1)\kappa}B^{p-1}\eta_\kappa^{(h)}.$$

At least the term in $B^{p-1}\eta_\kappa^{(p-1)}$ with coefficient $(-1)^{p-1}e^{-(p-1)\kappa}$ does not vanish. On the other hand, neither the term with lag zero in $\eta_\kappa^{(p-1)}$ vanishes, because its coefficient is $\alpha_{0,0}=1$.

General case. We now join the previous results for the general case with parameter $\boldsymbol{\kappa}$, a p-vector with p_j components equal to κ_j, $j=1,2,\ldots,q$, with $\kappa_1,\ldots,\kappa_q$ all different of each other and $\sum_{j=1}^q p_j = p$. We use the result of Theorem 1(1) and conclude that $x=\mathcal{OU}_{\boldsymbol{\kappa}}(\Lambda)$ has the same second-order moments as the ARMA$(p,p-1)$ model

$$\prod_{j=1}^{q}(1-\mathrm{e}^{-\kappa_j}B)^{p_j}x = \sum_{j=1}^{q}K_j\prod_{l\neq j}(1-\mathrm{e}^{-\kappa_l}B)^{p_l}\mathrm{MA}_j \tag{46}$$

with MA_j the moving average of order p_j-1 given by Equation (45).