Supplementary Material: Asymptotic properties of a componentwise ARH(1) plug-in predictor

J. Álvarez-Liébana †, D. Bosq ‡and M. D. Ruiz-Medina †

†*Department of Statistics and Operation Research (javialvaliebana@ugr.es, mruiz@ugr.es)*
Faculty of Sciences, University of Granada
Campus Fuente Nueva s/n
18071 Granada, Spain
‡*LSTA, Université Pierre et Marie Curie-Paris 6 (denis.bosq@upmc.fr)*
4 place Jussieu, 75005, Paris, France

Abstract

This document provides as supplementary material to the paper entitled *Asymptotic properties of a componentwise ARH(1) plug-in predictor*, a numerical example where the methodology proposed in such paper still works beyond the considered **Assumption A2**.

Non-diagonal autocorrelation operator

This section illustrates the performance of the proposed estimation methodology, when **Assumption A2** is not satisfied, but ρ is close to be diagonal in some sense. The numerical results obtained are compared to those ones derived from the computation of the ARH(1) predictors, based on the componentwise estimators proposed in [3, 4], where this diagonal assumption is not required. The Gaussian ARH(1) process generated has autocorrelation operator ρ, with coefficients $\rho_{j,h}$ with respect to the basis $\{\phi_j \otimes \phi_h\}_{j,\ h \geq 1}$, given by

$$\rho_{j,j}^2 = \left(\frac{\lambda_j \left((-\Delta)_{(a,b)} \right)}{\lambda_1 \left((-\Delta)_{(a,b)} \right) - \epsilon} \right)^{-\delta_2},$$

in the diagonal, and outside of the diagonal

$$\rho_{j,j+a}^2 = \frac{0.01}{5a^2},\ a = 1, \dots, 5, \quad \rho_{j+a,j}^2 = \frac{0.02}{5a^2},\ a = 1, \dots, 5, \tag{1}$$

Preprint submitted to Journal of Multivariate Analysis *October 7, 2016*

where $\rho^2_{j,j+a} = \rho^2_{j+a,j} = 0$ when $a \geq 6$. The coefficients of the autocovariance operator R_ε of the innovation process ε, with respect to the basis $\{\phi_j \otimes \phi_h\}_{j,h\geq 1}$ are given by $\sigma^2_{j,j} = C_j\left(1-\rho^2_{j,j}\right)$, in the diagonal, and outside of the diagonal by

$$\sigma^2_{j,j+a} = \frac{0.015}{5a^2},\ a=1,\ldots,5, \quad \sigma^2_{j+a,j} = \frac{0.01}{5a^2},\ a=1,\ldots,5, \tag{2}$$

where $\sigma^2_{j,j+a} = \sigma^2_{j+a,j} = 0$ when $a \geq 6$. Table 1 below displays the empirical truncated values of $\mathrm{E}\left\{\|\rho(X_{n-1}) - \widehat{\rho}_{k_n}(X_{n-1})\|_H\right\}$, based on $N = 200$ simulations of each one of the 20 functional samples considered, with sizes $n_t = 15000 + 20000(t-1)$, $t = 1,\ldots,20$, for the corresponding k_n values obtained, in this case, by the rule $k_n = n^{1/\alpha}$, with $\alpha = 6$. As in the paper, we have considered parameter $\delta_1 = 2.4$, in the definition of the eigenvalues of C, but, in this case, as noted before, operators ρ and R_ε are non-diagonal (see equations (1) and (2)). The estimators of ρ, and the associated plug-in predictors are computed, for the three approaches compared, under the assumption that the eigenvectors of C are known. Specifically, the following estimators and predictors are computed:

$$\widehat{\rho}_{k_n} = \sum_{j=1}^{k_n} \widehat{\rho}_{n,j}\phi_j \otimes \phi_j, \tag{3}$$

where, for each $j \geq 1$,

$$\widehat{\rho}_{n,j} = \frac{\widehat{D}_{n,j}}{\widehat{C}_{n,j}} = \frac{\frac{1}{n-1}\sum_{i=0}^{n-2} X_{i,j}X_{i+1,j}}{\frac{1}{n}\sum_{i=0}^{n-1} X^2_{i,j}} = \frac{n}{n-1}\frac{\sum_{i=0}^{n-2} X_{i,j}X_{i+1,j}}{\sum_{i=0}^{n-1} X^2_{i,j}}. \tag{4}$$

$$\widehat{\rho}_n(x) = \left(\Pi^{k_n} D_n \widehat{C}_n^{-1}\Pi^{k_n}\right)(x) = \sum_{l=1}^{k_n} \widehat{\rho}_{n,l}(x)\phi_l, \quad x \in H, \tag{5}$$

$$\widehat{\rho}_{n,l}(x) = \frac{1}{n-1}\sum_{i=0}^{n-2}\sum_{j=1}^{k_n} \frac{1}{\widehat{C}_{n,j}}\langle \phi_j, x\rangle_H X_{i,j}X_{i+1,l}, \tag{6}$$

$$\widehat{\rho}_{n,a}(x) = \left(\Pi^{k_n} D_n \widehat{C}_{n,a}^{-1}\Pi^{k_n}\right)(x) = \sum_{l=1}^{k_n} \widehat{\rho}_{n,a,l}(x)\phi_l, \quad x \in H, \tag{7}$$

$$\widehat{\rho}_{n,a,l}(x) \quad = \quad \frac{1}{n-1}\sum_{i=1}^{n-1}\sum_{j=1}^{k_n}\frac{1}{\max\left(\widehat{C}_{n,j},\ a_n\right)}\langle \phi_j, x\rangle_H X_{i,j}X_{i+1,l}, \tag{8}$$

As expected, in Table 1, an outperformance of the approaches in [3, 4] is observed in comparison with our methodology. However, for large sample sizes, the ARH(1) prediction methodology proposed here still can be applied with an order of magnitude of 10^{-2}, for the empirical errors associated with $\widehat{\rho}_{k_n}$, given in equations (3)–(4). Thus, in the *pseudodiagonal* autocorrelation operator case, in some sense, our approach could still be considered. As referred in our paper, an example is given in the case where the autocovariance and autocorrelation operators admit a sparse representation in terms of a suitable orthonormal wavelet basis (see, for instance, [1, 2]).

References

[1] C. Angelini, D. De Canditiis and F. Leblanc, Wavelet regression estimation in nonparametric mixed effect models, *J. Multivariate Anal.* 85, 2003, pp. 267–291.

[2] A. Antoniadis and T. Sapatinas, Wavelet methods for continuous-time prediction using Hilbert-valued autoregressive processes, *J. Multivariate Anal.* 87, 2003, pp. 133–158.

[3] D. Bosq, *Linear Processes in Function Spaces*, Springer-Verlag, New York 2000.

[4] S. Guillas, Rates of convergence of autocorrelation estimates for autoregressive Hilbertian processes, *Statistics & Probability Letters* 55, 2001, pp. 281–291.

n	k_n	Our approach	Bosq (2000)	Guillas (2001)
$n_1 = 15000$	4	0.5812	$8.94\,(10)^{-2}$	0.1055
$n_2 = 35000$	5	0.5604	$7.05\,(10)^{-2}$	$9.49\,(10)^{-2}$
$n_3 = 55000$	6	0.5480	$6.67\,(10)^{-2}$	$9.14\,(10)^{-2}$
$n_4 = 75000$	6	0.5322	$6.24\,(10)^{-2}$	$8.85\,(10)^{-2}$
$n_5 = 95000$	6	0.5115	$5.89\,(10)^{-2}$	$8.47\,(10)^{-2}$
$n_6 = 115000$	6	0.4975	$5.62\,(10)^{-2}$	$8.04\,(10)^{-2}$
$n_7 = 135000$	7	0.4946	$5.57\,(10)^{-2}$	$7.66\,(10)^{-2}$
$n_8 = 155000$	7	0.4810	$5.28\,(10)^{-2}$	$7.24\,(10)^{-2}$
$n_9 = 175000$	7	0.4735	$5.01\,(10)^{-2}$	$6.78\,(10)^{-2}$
$n_{10} = 195000$	7	0.4608	$4.90\,(10)^{-2}$	$6.30\,(10)^{-2}$
$n_{11} = 215000$	7	0.4424	$4.69\,(10)^{-2}$	$6.07\,(10)^{-2}$
$n_{12} = 235000$	7	0.4250	$4.45\,(10)^{-2}$	$5.82\,(10)^{-2}$
$n_{13} = 255000$	7	0.4106	$4.25\,(10)^{-2}$	$5.54\,(10)^{-2}$
$n_{14} = 275000$	8	0.4080	$4.14\,(10)^{-2}$	$5.16\,(10)^{-2}$
$n_{15} = 295000$	8	0.3808	$4.09\,(10)^{-2}$	$4.81\,(10)^{-2}$
$n_{16} = 315000$	8	0.3604	$3.85\,(10)^{-2}$	$4.53\,(10)^{-2}$
$n_{17} = 335000$	8	0.3489	$3.56\,(10)^{-2}$	$4.29\,(10)^{-2}$
$n_{18} = 355000$	8	0.3302	$3.29\,(10)^{-2}$	$3.98\,(10)^{-2}$
$n_{19} = 375000$	8	0.3204	$2.90\,(10)^{-2}$	$3.75\,(10)^{-2}$
$n_{20} = 395000$	8	0.3177	$2.62\,(10)^{-2}$	$3.44\,(10)^{-2}$

Table 1: *Non-diagonal case.* Truncated empirical values of $\mathrm{E}\left\{\|\rho(X_{n-1}) - \widehat{\rho}_{k_n}(X_{n-1})\|_H\right\}$, given in equations (3)–(4) (third column), in equations (5)–(6) (fourth column), and in equations (7)–(8) (fifth column), with $\delta_1 = 2.4$ and $\delta_2 = 1.1$, considering the sample sizes $n_t = 15000 + 20000(t-1)$, $t = 1, \dots, 20$, and the corresponding $k_n = n^{1/\alpha}$ values ($\alpha = 6$). The eigenvectors $\{\phi_j\}_{j\geq 1}$ are assumed to be known.