A WIDELY INFINITE PAST PREDICTION PROCEDURE

Jesús Navarro-Moreno, Rosa M. Fernández-Alcalá, Juan C. Ruiz-Molina and Antonia Oya
Department of Statistics and Operations Research, University of Jaén, 23071 Jaén, Spain

Keywords: Prediction theory, Transfer function models, Widely linear processing.

Abstract: Using a widely linear (WL) processing, a prediction algorithm has been designed for WL transfer function models in terms of an infinite number of past observations. This algorithm leads to a suboptimum infinite past predictor which approximates the optimal predictor based on a finite past information when the size of the series goes to infinite. Hence, the applicability of our solution lies in those situations where the predictor based on a finite past is difficult to obtain.

1 INTRODUCTION

Prediction based upon an infinite number of past observations is a problem of great relevance in statistical communication theory. Specifically, in those situations where the predictor based on finite past is difficult to obtain because of the number of available observations is extremely large, infinite past prediction problems provide feasible recursive algorithms for the computation of a suboptimum estimate which approximates the finite past predictor optimally [see, for example, (Brockwell and Davis, 1991)].

In particular, this strategy has been widely used in transfer function models (Box and Jenkins, 1970; Brockwell and Davis, 1991). Transfer function models, also called dynamic regression models, are extensions of familiar linear regression models which include not only information related to the past of the time series of interest but also the present and past values of other time series. Thus, the prediction of the first time series may be considerably improved by using information coming from the second.

On the other hand, the widely linear (WL) processing has provided a new perspective for solving several problems concerned with noncircular or improper complex-valued time series. This approach, based on the information supplied by both the signal and its conjugate, has shown its efficiency against the conventional or strictly linear (SL) processing in many areas of statistical signal processing such as modeling and estimation, among others [see, e.g., (Mandic and Goh, 2009; Navarro-Moreno, 2008; Navarro-Moreno et al., 2009; Picinbono and Chevalier, 1995; Picinbono and Bondon, 1997)]. Indeed, in the modeling field, WL systems appear to be more suitable than SL systems in the representation of this type of signal. In this framework, the WL finite past prediction problem for WL ARMA models has been studied in (Navarro-Moreno, 2008).

This paper tackles the WL infinite past prediction problem for a more general WL system than the one considered in (Navarro-Moreno, 2008). Specifically, the time series of interest is assumed to be modeled by a WL transfer function system and thus, following a WL processing, a recursive prediction algorithm is devised from the infinite past information supplied by both the input and output of such a model. This algorithm becomes an alternative approach to the WL finite past prediction problem of this type of system which, in general, is difficult to address. For this purpose, we first introduce WL transfer function models in Section 2. Next, the WL infinite past prediction problem is addressed in Section 3. Finally, an illustrative example is developed in Section 4.

2 WL TRANSFER FUNCTION MODELS

To start with, we introduce some important notations that will be used throughout the paper.

The real part of a complex number will be denoted by $\Re\{\cdot\}$, the transpose of a vector by $(\cdot)'$, the complex conjugate by $(\cdot)^*$ and the conjugate transpose by $(\cdot)^{\mathsf{H}}$. In general, we will consider the augmented version $\mathbf{X}_t = [X_t, X_t^*]'$ of the complex-valued random process X_t.

Navarro-Moreno J., M. Fernández-Alcalá R., C. Ruiz-Molina J. and Oya A..
A WIDELY INFINITE PAST PREDICTION PROCEDURE.
DOI: 10.5220/0003544704410444
In *Proceedings of the 8th International Conference on Informatics in Control, Automation and Robotics* (ICINCO-2011), pages 441-444
ISBN: 978-989-8425-74-4

Moreover, $\mathcal{M}_2$ represents the set of 2×2 complex-valued matrices[1]

$$\mathbf{A}_j = \begin{pmatrix} a_{1j} & a_{2j} \\ a_{2j}^* & a_{1j}^* \end{pmatrix}$$

Now, to attain the WL infinite past prediction problem, we consider a more general WL model than those previously suggested in (Picinbono and Bondon, 1997; Navarro-Moreno, 2008; Box and Jenkins, 1970). Specifically, let a process Y_t which is the output of the transfer function model

$$Y_t = \sum_{j=0}^{\infty} \left(t_{1j}X_{t-j} + t_{2j}X_{t-j}^*\right) + N_t \qquad (1)$$

which satisfies the following characteristics:

- The input process X_t satisfies the WL ARMA system (Navarro-Moreno, 2008)

$$X_t = \sum_{j=1}^{p_1} \left(g_{1j}X_{t-j} + g_{2j}X_{t-j}^*\right) + \sum_{j=0}^{q_1} \left(h_{1j}Z_{t-j} + h_{2j}Z_{t-j}^*\right) \qquad (2)$$

  where Z_t is a centered doubly white noise with correlation function $E[Z_iZ_j^*] = d_1\delta(i-j)$ and complementary function $E[Z_iZ_j] = d_2\delta_{ij}$, with $|d_2| < d_1$ and δ_{ij} the Kronecker delta function.

- The noise N_t is supposed to be generated by a WL system

$$N_t = \sum_{j=1}^{p_2} \left(m_{1j}N_{t-j} + m_{2j}N_{t-j}^*\right) + \sum_{j=0}^{q_2} \left(l_{1j}W_{t-j} + l_{2j}W_{t-j}^*\right) \qquad (3)$$

  with $E[W_iW_j^*] = e_1\delta(i-j)$, $E[W_iW_j] = e_2\delta(i-j)$, $|e_2| < e_1$ and the augmented noises $\mathbf{Z}_t$ and $\mathbf{W}_t$ are uncorrelated.

As it is usual in the prediction process for transfer function models, the predictor of Y_t based on a finite past is, in general, difficult to obtain and the only simple way to compute the predictor is by using the infinite past (Box and Jenkins, 1970; Brockwell and Davis, 1991). Thus, our aim here is to predict the process Y_{n+s} based on the infinite joint past $\{[Y_t, X_t]', -\infty < t \leq n\}$ under a WL processing. Specifically, expressions for computing this WL infinite past predictor, denoted by $\hat{Y}_{n+s}^{WL}$, as well as its associated mean square error are provided in the next section. The proofs and further details about these results here can be found in (Navarro-Moreno et al., 2011).

[1] $\mathcal{M}_2$ constitutes a matrix algebra which is closed under addition, multiplication, inversion (when inverses exist), and multiplication with a real, but not with a complex scalar.

3 WL INFINITE PAST PREDICTION

First of all, we must note that the WL infinite past predictor $\hat{Y}_{n+s}^{WL}$ is the projection of $\mathbf{Y}_{n+s}$ onto the space[2] $\mathcal{H}_\infty = \overline{\text{sp}}\{[\mathbf{Y}_t, \mathbf{X}_t]', -\infty < t \leq n\}$.

Then, introducing the following three types of matrix operators

$$\mathbf{G}_p^-(B) := \mathbf{I} - \sum_{i=1}^{p} \mathbf{G}_i B^i$$
$$\mathbf{H}_q(B) := \sum_{j=0}^{q} \mathbf{H}_j B^j$$
$$\mathbf{T}(B) := \sum_{k=0}^{\infty} \mathbf{T}_k B^k$$

with $\mathbf{I}$ the identity matrix, B^j the backward shift operator ($B^jX_t = X_{t-j}$) and $\mathbf{G}_i, \mathbf{H}_j, \mathbf{T}_k \in \mathcal{M}_2$, $i = 1,\ldots,p$, $j = 0,1,\ldots,q$, $k = 0,1,\ldots$, equations (1), (2) and (3) can be rewritten in terms of the augmented processes $\mathbf{Y}_t$, $\mathbf{X}_t$, $\mathbf{Z}_t$, $\mathbf{N}_t$ and $\mathbf{W}_t$ as

$$\mathbf{Y}_t = \mathbf{T}(B)\mathbf{X}_t + \mathbf{N}_t$$
$$\mathbf{G}_{p_1}^-(B)\mathbf{X}_t = \mathbf{H}_{q_1}(B)\mathbf{Z}_t$$
$$\mathbf{M}_{p_2}^-(B)\mathbf{N}_t = \mathbf{L}_{q_2}(B)\mathbf{W}_t$$

and hence, it is clear that $\mathcal{H}_\infty = \overline{\text{sp}}\{[\mathbf{Z}_t, \mathbf{W}_t]', -\infty < t \leq n\}$. This fact leads to the following expressions for the WL infinite past predictor $\mathbf{Y}_{n+s}$ as well as its mean square error

Theorem 1. *The WL infinite past predictor $\hat{Y}_{n+s}^{WL}$ of the process Y_t given by* (1), *has the following form*

$$\hat{Y}_{n+s}^{WL} = \sum_{j=s}^{\infty} \left(a_{1j}Z_{n+s-j} + a_{2j}Z_{n+s-j}^*\right) + \sum_{j=s}^{\infty} \left(f_{1j}W_{n+s-j} + f_{2j}W_{n+s-j}^*\right) \qquad (4)$$

where the coefficients a_{1j}, a_{2j}, f_{1j}, f_{2j} are obtained from the equations

$$\sum_{j=0}^{\infty} \mathbf{A}_j B^j = \mathbf{T}(B)(\mathbf{G}_{p_1}^-)^{-1}(B)\mathbf{H}_{q_1}(B)$$
$$\sum_{j=0}^{\infty} \mathbf{F}_j B^j = (\mathbf{M}_{p_2}^-)^{-1}(B)\mathbf{L}_{q_2}(B) \qquad (5)$$

[2] $\overline{\text{sp}}\{[\mathbf{Y}_t, \mathbf{X}_t]', -\infty < t \leq n\}$ denotes the closed span of the vectors set $\{[\mathbf{Y}_t, \mathbf{X}_t]', -\infty < t \leq n\}$.

with $\mathbf{A}_j, \mathbf{F}_j \in \mathcal{M}_2$, $j = 0, 1, \ldots$ *Also the error of* $\hat{Y}^{WL}_{n+s}$ *is*

$$\begin{aligned}\mathrm{Error}(\hat{Y}^{WL}_{n+s}) &= E\left[\left|Y_{n+s} - \hat{Y}^{WL}_{n+s}\right|^2\right] \\ &= \sum_{j=0}^{s-1}\left(2\Re\{a_{1j}d_2a^*_{2j}\} + a_{1j}d_1a^*_{1j} + a_{2j}d_1a^*_{2j}\right) \\ &+ \sum_{j=0}^{s-1}\left(2\Re\{f_{1j}e_2f^*_{2j}\} + f_{1j}e_1f^*_{1j} + f_{2j}e_1f^*_{2j}\right) \quad (6)\end{aligned}$$

Representation (4) is not convenient from the computational point of view since it depends on an infinite number of past observations and thus another expression is necessary. For this purpose, using operators of the form (Box and Jenkins, 1970)

$$\mathbf{T}(B) := (\mathbf{V}^-_{p_3})^{-1}(B)\mathbf{R}_{q_3}(B)B^b$$

with $\mathbf{V}_i, \mathbf{R}_j \in \mathcal{M}_2$, $i = 1, \ldots, p_3$ and $j = 0, 1, \ldots, q_3$, a recursive expression for computing the WL infinite past predictor $\mathbf{Y}_{n+s}$ is derived in next theorem

Theorem 2. *The WL infinite past predictor* $\hat{Y}^{WL}_{n+s}$ *of the process* Y_t *given by* (1), *can be computed as follows:*

$$\begin{aligned}\hat{Y}^{WL}_{n+s} &= \sum_{j=1}^{p_3}\left(v_{1j}\hat{Y}^{WL}_{n+s-j} + v_{2j}\hat{Y}^{WL^*}_{n+s-j}\right) \\ &+ \sum_{j=0}^{q_3}\left(r_{1j}\hat{X}^{WL}_{n+s-b-j} + r_{2j}\hat{X}^{WL^*}_{n+s-b-j}\right) \\ &+ \sum_{j=s}^{p_4}\left(c_{1j}W_{n+s-j} + c_{2j}W^*_{n+s-j}\right) \quad (7)\end{aligned}$$

with $\hat{Y}^{WL}_j = Y_j$ *and* $\hat{X}^{WL}_j = X_j$, $j = 1, \ldots, n$ *and where* $\hat{X}^{WL}_j$ *is the WL predictor of* X_j *calculated throughout the expressions*

$$\begin{aligned}\hat{X}^{WL}_{n+1} &= -\sum_{j=1}^{\infty}\left(\bar{k}_{1j}X_{n+1-j} + \bar{k}_{2j}X^*_{n+1-j}\right) \\ \hat{X}^{WL}_{n+2} &= -\bar{k}_{1,1}\hat{X}^{WL}_{n+1} - \bar{k}_{2,1}\hat{X}^{WL^*}_{n+1} \\ &\quad - \sum_{j=2}^{\infty}\left(\bar{k}_{1j}X_{n+2-j} + \bar{k}_{2j}X^*_{n+2-j}\right) \quad (8) \\ &\vdots\end{aligned}$$

Moreover, for $s \leq p_4$, *the coefficients* c_{1j}, c_{2j} *are the elements of* $\mathbf{C}_j$, *obtained from the equation*

$$\sum_{j=0}^{p_4}\mathbf{C}_jB^j = \mathbf{V}^-_{p_3}(B)(\mathbf{M}^-_{p_2})^{-1}(B)\mathbf{L}_{q_2}(B)$$

with $\mathbf{C}_j \in \mathcal{M}_2$, $j = 0, \ldots, p_4$ *and, for* $s > p_4$, *the last term in* (7) *vanishes.*

Remark 1. *For large n, we can define a WL suboptimum predictor by truncating* (8) *at n terms and replacing in* (7), *the predictors* $\hat{X}^{WL}_j$ *by the approximate predictors* $\tilde{X}^{WL}_j$ *given by the expressions*

$$\begin{aligned}\tilde{X}^{WL}_{n+1} &= -\sum_{j=1}^{n}\left(\bar{k}_{1j}X_{n+1-j} + \bar{k}_{2j}X^*_{n+1-j}\right) \\ \tilde{X}^{WL}_{n+2} &= -\bar{k}_{1,1}\tilde{X}^{WL}_{n+1} - \bar{k}_{2,1}\tilde{X}^{WL^*}_{n+1} \\ &\quad - \sum_{j=2}^{n+1}\left(\bar{k}_{1j}X_{n+2-j} + \bar{k}_{2j}X^*_{n+2-j}\right) \\ &\vdots \\ \tilde{X}^{WL}_{n+s} &= -\sum_{j=1}^{n+s-1}\left(\bar{k}_{1j}\tilde{X}^{WL}_{n+s-j} + \bar{k}_{2j}\tilde{X}^{WL^*}_{n+s-j}\right)\end{aligned}$$

with $\tilde{X}^{WL}_j = X_j$, $j = 1, \ldots, n$.

The performance of the resultant finite past predictor can be assessed by comparing its error with the lower bound found in (6).

4 NUMERICAL EXAMPLE

Consider the WL transfer function model

$$Y_t = X_t + \exp\{5\mathrm{j}\}X^*_t + N_t$$

where $\mathrm{j} = \sqrt{-1}$ and X_t and N_t are the following WL MA(1) and MA(2) models respectively

$$\begin{aligned}X_t &= Z_t + Z_{t-1} \\ N_t &= W_t + 0.5W_{t-1} + 2W^*_{t-1} + 3W^*_{t-2}\end{aligned}$$

with $E[Z_iZ^*_j] = \delta(i-j)$, $E[Z_iZ_j] = d_2\delta(i-j)$, $E[W_iW^*_j] = \delta(i-j)$ and $E[W_iW_j] = e_2\delta(i-j)$.

We carry out an analysis of prediction for $s = 1, 2$ in function of d_2 and e_2, with d_2 and e_2 varying between 0 and 0.99. Denote the errors associated with the WL and with SL predictors for every value d_2 and e_2 by $\mathrm{Error}(\hat{Y}^{WL}_{n+s}(d_2,e_2))$ and $\mathrm{Error}(\hat{Y}^{SL}_{n+s}(d_2,e_2))$, respectively. From (6) it can be shown that

$$\begin{aligned}\mathrm{Error}(\hat{Y}^{WL}_{n+1}(d_2,e_2)) &= 2\Re\{\exp\{-5\mathrm{j}\}d_2\} + 3 \\ \mathrm{Error}(\hat{Y}^{WL}_{n+2}(d_2,e_2)) &= 4\Re\{\exp\{-5\mathrm{j}\}d_2\} + 2e_2 + 9.25\end{aligned}$$

Figures 1 and 2 depict the following error differences: $\mathrm{Error}(\hat{Y}^{SL}_{n+1}(d_2,e_2)) - \mathrm{Error}(\hat{Y}^{WL}_{n+1}(d_2,e_2))$ and $\mathrm{Error}(Y^{SL}_{n+2}(d_2,e_2)) - \mathrm{Error}(Y^{WL}_{n+2}(d_2,e_2))$, respectively. We can observe that the WL predictor has a slight better performance in the case of one-stage prediction than in the case of two-stage prediction, that is, the WL one-ahead predictor attains a greater difference with respect to the SL one-ahead predictor than that achieved by the WL two-ahead predictor in relation to the SL. Moreover, the noise N_t has a greater influence on the difference of errors than the noise Z_t, i.e., we observe a more significant change in this difference if a value of d_2 is fixed and we vary e_2 than if we fix a value of e_2 while d_2 varies. Finally, the advantages of WL processing are lost when $s > 2$ since the WL and the SL predictors coincide.

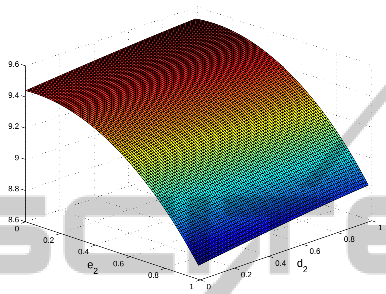

Figure 1: Error$(\hat{Y}^{SL}_{n+1}(d_2,e_2))$ – Error$(\hat{Y}^{WL}_{n+1}(d_2,e_2))$.

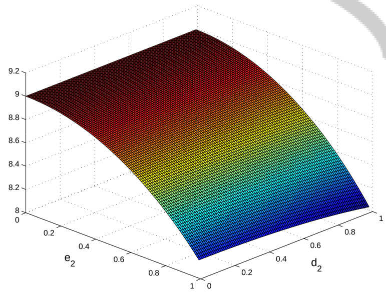

Figure 2: Error$(\hat{Y}^{SL}_{n+2}(d_2,e_2))$ – Error$(\hat{Y}^{WL}_{n+2}(d_2,e_2))$.

ACKNOWLEDGEMENTS

This work was supported in part by Project MTM2007-66791 of the Plan Nacional de I+D+I, Ministerio de Educación y Ciencia, Spain, which is financed jointly by the FEDER.

REFERENCES

Box, G. and Jenkins, G. (1970). *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.

Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods*. Springer-Verlag, New York, 2nd edition.

Mandic, D. P. and Goh, V. S. L. (2009). *Complex Valued Nonlinear Adaptative Filters. Noncircularity, Widely Linear and Neural Models*. Wiley.

Navarro-Moreno, J. (2008). ARMA Prediction of Widely Linear Systems by Using the Innovations Algorithm. *IEEE, Trans. Signal Processing*, 56(7):3061–3068.

Navarro-Moreno, J., Moreno-Kayser, J., Fernández-Alcalá, R. M., and Ruiz-Molina, J. C. (2009). Widely Linear Estimation Algorithms for Second-Order Stationary Signals. *IEEE, Trans. Signal Processing*, 51(1):306–312.

Navarro-Moreno, J., Moreno-Kayser, J., Fernández-Alcalá, R. M., and Ruiz-Molina, J. C. (2011). Widely Linear Prediction for Transfer Function Models Based on the Infinite Past. *Computational Statistics & Data Analysis*. In press.

Picinbono, B. and Bondon, P. (1997). Second-Order Statistics of Complex Signals. *IEEE, Trans. Signal Processing*, 45(2):411–420.

Picinbono, B. and Chevalier, P. (1995). Widely Linear Estimation with Complex Data. *IEEE, Trans. Signal Processing*, 43(8):2030–2033.