RECURSIVE METHODS FOR ESTIMATING MULTIPLE MISSING VALUES OF A MULTIVARIATE STATIONARY PROCESS

Pascal Bondon

Laboratoire des Signaux et Systèmes,
CNRS/ESE/Université de Paris-Sud
Plateau de Moulon,
91192 Gif-sur-Yvette, France

Diego P. Ruiz and Antolino Gallego

Departamento de Física Aplicada,
Facultad de Ciencias,
Universidad de Granada,
18071 Granada, Spain

ABSTRACT

Existing methods for estimating linearly s future values of a m-variate stationary random process using a record of p vectors from the past consist in first solving the one-step prediction problem and then all the h-step prediction problems for $2 \leq h \leq s$ independently. When the Levinson algorithm is used, each prediction problem is solved with a numerical complexity proportional to p^2. In this paper, we propose new methods to solve the h-step prediction problems for $h \geq 2$ with a numerical complexity proportional to p.

1. INTRODUCTION

In many signal processing problems arising for example in geophysics, communications, and neurophysics, as well as in statistical time series analysis [3], it is a major concern to develop a model of the underlying data series. When the model is m-variate autoregressive linear and the theory of linear prediction of stationary random vectors is used to calculate the model, a system of linear equations must be solved [6]. The direct solution of this system requires $(mp)^3$ multiplication and divisions, p being the order of the predictor. The use of the multivariate Levinson-Durbin algorithm [5, 7, 8] allows a large computational saving since it only requires m^3p^2 operations.

In many instances, several observations of the m-variate process are missing or erroneous. In these cases, it may be of interest, not only to solve the one-step linear prediction problem, but also some h-step prediction problems for $h \geq 1$ [1, 3, 4]. Each h-step prediction problem can be solved separately with the multivariate Levinson algorithm. Since this algorithm is only order recursive and not step recursive, its complexity increases linearly with the final step s. This fact may be a severe drawback for large values of s. It is thus of interest to investigate methods which do not suffer from this limitation. The idea is to find an algorithm which is not only recursive on the order of the predictor, like the Levinson algorithm, but which also makes use of some recursivity on the step h.

In this paper we focus on the h-step ahead linear prediction problem and new relationships among the predictors are demonstrated. Based on these relations, order and step recursive algorithms with a reduced computational cost are proposed. Our algorithms solve the h-step prediction problems for $h \geq 2$ with a numerical complexity proportional to m^3p instead of m^3p^2.

2. PRELIMINARIES

Let $(x_k)_{k\in\mathbb{Z}}$ be a zero-mean m-variate stationary real random process defined on a probability space $(\Omega, \mathcal{F}, \mathsf{P})$. The Hilbert space of square-integrable univariate random variables is denoted by $\mathcal{L}^2 = \mathcal{L}^2(\Omega, \mathcal{F}, \mathsf{P})$. The Hilbert space of square-integrable m-variate random vectors is the product space $(\mathcal{L}^2)^m$. Its scalar product is defined by $\langle X \mid Y\rangle = \mathsf{E}(X^TY)$ and its norm by $\|X\| = \mathsf{E}(X^TX)^{1/2}$. For any integer k, x_k may be written as $x_k = (x_{k,1}, \dots, x_{k,m})^T$. Given a final step s, the problem is to find for each step h, $1 \leq h \leq s$, the linear estimate $\widehat{x}_{p+h}$ of x_{p+h} from the random vectors $(x_i)_{h\leq i\leq p}$. For any h, $(A_{h,i})_{1\leq i\leq p-h+1}$ denote the matrices such that

$$\widehat{x}_{p+h} = \sum_{i=1}^{p-h+1} A_{h,i}\, x_{p+1-i} \tag{1}$$

and $\Sigma_h = \mathsf{E}\left[(x_{p+h} - \widehat{x}_{p+h})(x_{p+h} - \widehat{x}_{p+h})^T\right]$ denotes the h-step prediction error matrix. The problem is therefore equivalent to finding the matrices $(A_{h,i})_{1\leq i\leq p-h+1}$ and the error matrices Σ_h for $1 \leq h \leq s$. Matrices $(A_{h,i})_{1\leq i\leq p-h+1}$ are determined by the orthogonality relations

$$\langle\!\langle x_{p+h} - \widehat{x}_{p+h} \mid x_{p+1-j}\rangle\!\rangle = 0 \tag{2}$$

for $j = 1, \dots, p-h+1$, where $\langle\!\langle X \mid Y\rangle\!\rangle = \mathsf{E}(XY^T)$. Let $(\Gamma_\tau)_{\tau\in\mathbb{Z}}$ be the correlation function of $(x_k)_{k\in\mathbb{Z}}$ given by the matrix $\Gamma_\tau = \mathsf{E}(x_k\, x_{k-\tau}^T)$. Using (1), (2) is equivalent to

$$\sum_{i=1}^{p-h+1} A_{h,i}\, \Gamma_{j-i} = \Gamma_{h-1+j} \tag{3}$$

for $j = 1, \dots, p-h+1$. When $h = 1$, we obtain the one-step linear multivariate prediction problem. In this case, it is well known that (3) can be solved using the multivariate Levinson-Durbin algorithm given in [8]. For ease of notation, $\mathrm{sp}(x)$ (where $x \in (\mathcal{L}^2)^m$) denotes the product space $(\mathrm{sp}(x_j)_{1\le j\le m})^m$. Therefore, any element y of $\mathrm{sp}(x)$ can be written as $y = A\,x$, where $A \in \mathbb{R}^{(m\times m)}$. Let $\mathcal{H}$ be a subspace of $(\mathcal{L}^2)^m$. The definition of $\mathrm{sp}(x)$ implies that $\mathcal{H}$ and $\mathrm{sp}(x)$ are orthogonal if and only if $\langle u \mid A\,x\rangle = 0$ for all $u \in \mathcal{H}$ and for all $A \in \mathbb{R}^{(m\times m)}$. This is equivalent to $\langle\!\langle u \mid x\rangle\!\rangle = 0$ for all $u \in \mathcal{H}$. For any finite integers l, n satisfying $l \le n$, we denote $\mathcal{H}_{l,n} = \mathrm{sp}(x_i)_{l\le i\le n} = (\mathrm{sp}(x_{i,j})_{l\le i\le n, 1\le j\le m})^m$. For any step $h \ge 1$, and for any order $n \ge 1$, $\widehat{x}_{n+h}^h$ denotes the linear estimate of x_{n+h} based on $(x_i)_{1\le i\le n}$. $\widehat{x}_{n+h}^h$ is therefore the orthogonal projection of x_{n+h} onto $\mathcal{H}_{1,n}$, i.e. $\widehat{x}_{n+h}^h = \mathbb{P}[x_{n+h} \mid \mathcal{H}_{1,n}]$. The matrices $(\Phi_{n,i}^h)_{1\le i\le n}$ are such that

$$\widehat{x}_{n+h}^h = \sum_{i=1}^{n} \Phi_{n,i}^h\, x_{n+1-i} \tag{4}$$

and $V_n^h = \langle\!\langle x_{n+h} - \widehat{x}_{n+h}^h \mid x_{n+h} - \widehat{x}_{n+h}^h\rangle\!\rangle$ denotes the mean-square error matrix. Then, the stationarity of $(x_k)_{k\in\mathbb{Z}}$ yields $(\forall h \in \{1, \dots, s\})$, $(\forall i \in \{1, \dots, p-h+1\})$,

$$\begin{aligned} A_{h,i} &= \Phi_{p-h+1,i}^h \\ \Sigma_h &= V_{p-h+1}^h\,. \end{aligned} \tag{5}$$

In the following we establish purely order recursive, purely step recursive, and mixed order and step recursive relations between the h-step predictors and their respective error matrices. Each kind of relation is obtained by an adequate decomposition of the observation space $\mathcal{H}_{1,n}$. For clarity, the relations are given in different propositions. All the propositions presented in the paper are demonstrated in [2]. They hold under the hypothesis:

Hypothesis 1. The correlation matrix of vector X_n defined by $X_n = (x_1^T, \dots, x_n^T)^T$ is nonsingular for any integer $n \ge 1$.

This hypothesis is not restrictive at all, and allows us to avoid the singular case where the random process $(x_k)_{k\in\mathbb{Z}}$ is linearly deterministic with a finite past.

3. ORDER RECURSIVE ALGORITHM

When $h = 1$, (3) is solved using the multivariate Levinson-Durbin algorithm. This method gives all the predictors $\widehat{x}_{n+1}^1$ for $1 \le n \le p$. In contrast to the univariate algorithm, the multivariate version requires the solution of two sets of linear equations, one arising in the calculation of the forward predictor $\widehat{x}_{n+1}^1$, and the other in the calculation of the backward predictor $\mathbb{P}[x_0 \mid \mathcal{H}_{1,n}]$. This is due to the fact that matrix Γ_τ is generally not symmetric when $\tau \ne 0$. The same situation appears when dealing with the h-step prediction problem where $h > 1$. For any $h \ge 1$, $(\ddot{\Phi}_{n,i}^h)_{1\le i\le n}$ denotes the matrices such that

$$\mathbb{P}[x_{-h+1} \mid \mathcal{H}_{1,n}] = \sum_{i=1}^{n} \ddot{\Phi}_{n,i}^h\, x_i \tag{6}$$

and $\ddot{V}_n^h = \big\langle\!\big\langle x_{-h+1} - \mathbb{P}[x_{-h+1} \mid \mathcal{H}_{1,n}] \,\big|\, x_{-h+1} - \mathbb{P}[x_{-h+1} \mid \mathcal{H}_{1,n}]\big\rangle\!\big\rangle$ denotes the h-step backward prediction error matrix. In the following proposition, we give the order recursive expressions of the matrices $(\Phi_{n,i}^h)_{1\le i\le n}$, $(\ddot{\Phi}_{n,i}^h)_{1\le i\le n}$, V_n^h, and $\ddot{V}_n^h$.

Proposition 1. $(\forall h \ge 1)$, $(\forall n > 1)$, $(\forall i \in \{1, \dots, n-1\})$, we have $\Phi_{1,1}^h = \Gamma_h \Gamma_0^{-1}$, $\ddot{\Phi}_{1,1}^h = \Gamma_{-h}\Gamma_0^{-1}$, $V_1^h = \Gamma_0 - \Gamma_h \Gamma_0^{-1}\Gamma_h^T$, $\ddot{V}_1^h = \Gamma_0 - \Gamma_{-h}\Gamma_0^{-1}\Gamma_{-h}^T$, and

$$\begin{aligned} \Phi_{n,n}^h = {} & [\Gamma_{n+h-1} - \textstyle\sum_{i=1}^{n-1} \Gamma_{n+h-i-1}\, (\ddot{\Phi}_{n-1,i}^1)^T] \\ & \times (\ddot{V}_{n-1}^1)^{-1} \end{aligned} \tag{7}$$

$$\Phi_{n,i}^h = \Phi_{n-1,i}^h - \Phi_{n,n}^h\, \ddot{\Phi}_{n-1,n-i}^1 \tag{8}$$

$$\begin{aligned} \ddot{\Phi}_{n,n}^h = {} & [\Gamma_{-n-h+1} - \textstyle\sum_{i=1}^{n-1} \Gamma_{-n-h+i+1}\, (\Phi_{n-1,i}^1)^T] \\ & \times (V_{n-1}^1)^{-1} \end{aligned} \tag{9}$$

$$\ddot{\Phi}_{n,i}^h = \ddot{\Phi}_{n-1,i}^h - \ddot{\Phi}_{n,n}^h\, \Phi_{n-1,n-i}^1 \tag{10}$$

$$V_n^h = V_{n-1}^h - \Phi_{n,n}^h\, \ddot{V}_{n-1}^1\, (\Phi_{n,n}^h)^T \tag{11}$$

$$\ddot{V}_n^h = \ddot{V}_{n-1}^h - \ddot{\Phi}_{n,n}^h\, V_{n-1}^1\, (\ddot{\Phi}_{n,n}^h)^T. \tag{12}$$

When $h = 1$ in proposition 1, equations (7) and (9) can be respectively rewritten as

$$\Phi_{n,n}^1 = \left[\Gamma_n - \sum_{i=1}^{n-1} \Phi_{n-1,i}^1\, \Gamma_{n-i}\right] \left(\ddot{V}_{n-1}^1\right)^{-1}$$

$$\ddot{\Phi}_{n,n}^1 = \left[\Gamma_{-n} - \sum_{i=1}^{n-1} \ddot{\Phi}_{n-1,i}^1\, \Gamma_{-n+i}\right] \left(V_{n-1}^1\right)^{-1}$$

which give the multivariate Levinson-Durbin algorithm [3, pp. 422–423]. For each n, the relations given in proposition 1 involve $2m^3(2n+1) + 2q$ products where q stands for the number of multiplications to invert the $m \times m$ prediction error matrix V_{n-1}^1. Summing from $n = 1$ to $n = p$, we obtain $2m^3p^2 + O(p)$ multiplications, which is the numerical complexity to solve the one step prediction problem. Now when $h > 1$, the h-step prediction problem can be solved recursively on the order n up to $n = p - h + 1$ using the results for $h = 1$ and the relations (7), (8), and (11). For each

n and h, these relations involve $m^3(2n+1)+q$ products. Summing from $n=1$ to $n=p-h+1$, and taking into account that $h \ll p$ in practice, we obtain $m^3p^2+O(p)$ multiplications. Therefore, the multiple missing value estimation problem is solved with the complexity $m^3p^2(s-1)+O(p)$.

4. STEP RECURSIVE ALGORITHMS

We now present two step recursive algorithms for solving the h-step prediction problems for $h > 1$. For both methods, the one-step problem is solved using the Levinson-Durbin algorithm. The first algorithm uses the relations established in proposition 2. These relations may be interpreted as the equivalent of equations (8) and (11) used in the Levinson algorithm when the recursion is made on the step instead of on the order. More precisely the relation (13) [resp. (14)] in proposition 2 gives the expression of $\Phi^h_{n,i}$ [resp. V^h_n] in terms of $\Phi^{h-1}_{n,i+1}$ [resp. V^{h-1}_n] for a fixed n, whereas relation (8) [resp. (11)] in proposition 1 gives the expression of $\Phi^h_{n,i}$ [resp. V^h_n] in terms of $\Phi^h_{n-1,i}$ [resp. V^h_{n-1}] for a fixed h. The second algorithm uses the relations established in proposition 3. These relations give an expression of $\Phi^h_{n,i}$ in terms of the coefficients $\Phi^j_{n,i}$ for $1 \le j < h$.

Proposition 2. $(\forall h > 1), (\forall n > 1), (\forall i \in \{1, \dots, n-1\})$, we have

$$\Phi^h_{n,i} = \Phi^{h-1}_{n,i+1} + \Phi^{h-1}_{n,1}\,\Phi^1_{n-1,i} - \Phi^h_{n,n}\,\ddot{\Phi}^1_{n-1,n-i} \qquad (13)$$

$$\begin{aligned} V^h_n = {} & V^{h-1}_n + \Phi^{h-1}_{n,1}\,V^1_{n-1}\,(\Phi^{h-1}_{n,1})^T \\ & - \Phi^h_{n,n}\,\ddot{V}^1_{n-1}\,(\Phi^h_{n,n})^T. \end{aligned} \qquad (14)$$

Using the results of proposition 2, we propose the following algorithm to solve the h-step prediction problems for $1 < h \le s$. This method is referred to as the SR1 algorithm. Take $n = p-s+1$. Compute the matrices Φ^h_n and the error matrices V^h_n for all h, $2 \le h \le p+1-n$, using (7) for calculating $\Phi^h_{n,n}$, (13) for calculating the matrices $\Phi^h_{n,i}$ for $1 \le i \le n-1$, and (14) for calculating V^h_n. According to (5), for the last iteration $h = p+1-n$, we obtain the matrices $A_{p+1-n,i}$ for $i = 1, \dots, n$ and the error matrix Σ_{p+1-n}. Then, repeat the same procedure with $n = n+1$ until n reaches its highest value, $p-1$. The numerical complexity of the SR1 algorithm is $\frac{3}{2}m^3p(s^2-s)$.

Proposition 3. $(\forall h > 1), (\forall n \ge 1), (\forall i \in \{1, \dots, n\})$, we have

$$\Phi^h_{n,i} = \Phi^1_{n+h-1,i+h-1} + \sum_{j=1}^{h-1} \Phi^1_{n+h-1,h-j}\,\Phi^j_{n,i}. \qquad (15)$$

Proposition 3 gives a new algorithm, called the SR2 algorithm, which allows the h-step prediction problems to be solved for $1 < h \le s$. For each order n, $p-s+1 \le n \le p-1$, compute the matrices $\Phi^h_{n,i}$ for $i = 1, \dots, n$ and $2 \le h \le p+1-n$ using (15), and the error matrices V^h_n using (14). The numerical complexity of the SR2 algorithm is $\frac{1}{6}m^3p(s^3-s)$.

5. ORDER AND STEP RECURSIVE ALGORITHM

In this section, we present an algorithm which is order and step recursive for solving the h-step prediction problems for $h > 1$. The one-step problem is solved using the Levinson recursion. The algorithm uses the relations established in the following proposition:

Proposition 4. $(\forall h > 1), (\forall n \ge 1), (\forall i \in \{1, \dots, n\})$, we have

$$\Phi^h_{n,i} = \Phi^{h-1}_{n+1,i+1} + \Phi^{h-1}_{n+1,1}\,\Phi^1_{n,i} \qquad (16)$$

$$V^h_n = V^{h-1}_{n+1} + \Phi^{h-1}_{n+1,1}\,V^1_n\,\left(\Phi^{h-1}_{n+1,1}\right)^T. \qquad (17)$$

We now show that (16) allows the h-step prediction matrices $A_{h,i}$ for $1 \le h \le s$ and $i = 1, \dots, p-h+1$ to be calculated recursively, and that (17) is a step recursive relation for the error matrices Σ_h. Fix h, $1 < h \le s$. Taking $n = p-h+1$ in (16) and (17) and using (5), we obtain respectively

$$A_{h,i} = A_{h-1,i+1} + A_{h-1,1}\,\Phi^1_{p-h+1,i} \qquad (18)$$

$$\Sigma_h = \Sigma_{h-1} + A_{h-1,1}\,V^1_{p-h+1}\,\left(A_{h-1,1}\right)^T. \qquad (19)$$

In view of (18) and (19), the algorithm for estimating multiple missing values appears simply. This algorithm is referred to as the OSR algorithm. First, compute the prediction matrices $(\Phi^1_{n,i})_{1 \le i \le n}$ and the mean-square error matrices V^1_n for $1 \le n \le p$ using the multivariate Levinson algorithm. Next, compute the matrices $(A_{h,i})_{1 \le i \le p-h+1}$ and the prediction error matrices Σ_h for $2 \le h \le s$ using (18) and (19). The numerical complexity is $m^3p(s-1)$.

6. NUMERICAL SIMULATIONS

Numerical simulation results are now presented to compare the above algorithms. In all cases, the multiple missing value estimation problem is solved assuming that the correlation function $(\Gamma_\tau)_{\tau \in \mathbb{Z}}$ of $(x_k)_{k \in \mathbb{Z}}$ is known for $0 \le \tau \le p$. The numerical complexity in terms of number of multiplications of these algorithms to calculate the matrices $A_{h,i}$ for $1 \le i \le p-h+1$ and the error matrices Σ_h for $1 \le h \le s$ is summarized in Table 1. Since the one-step prediction problem is solved in each case using the classical Levinson algorithm, this table shows separately the numerical complexity for $h = 1$, and the complexity for $h = 2$ up to $h = s$. In the third column of Table 1, note that the Levinson algorithm has a complexity proportional to p^2, whereas the complexity of the three proposed recursive algorithms is proportional to p. Moreover, the differences among the three step recursive algorithms lie in a different proportionality factor for each method. So, the higher numerical complexity is for the SR2 and SR1 algorithms, which have a proportionality factor of s^3 and s^2, respectively. However, in the OSR algorithm this factor is s, and it is therefore the most suitable algorithm of those proposed in this paper. This result is due to the fact that the SR2 and SR1 algorithms are only step recursive, while the OSR algorithm is both step and order recursive.

	step one	h-steps for $2 \leq h \leq s$
Levinson	$2m^3p^2 + O(p)$	$m^3p^2(s-1) + O(p)$
SR1	$2m^3p^2 + O(p)$	$\frac{3}{2}m^3p(s^2 - s)$
SR2	$2m^3p^2 + O(p)$	$\frac{1}{6}m^3p(s^3 - s)$
OSR	$2m^3p^2 + O(p)$	$m^3p(s-1)$

Table 1: Numerical complexity of the algorithms.

Figure 1 plots the number of flops obtained when the predictors and the errors are calculated for different values of p in the range $[10, 500]$ and for $s = 5$. The number of flops means the floating point operations (one for each addition, product, and division). It is clear that the OSR algorithm provides a significant improvement over the Levinson algorithm and the improvement is an increasing function of p.

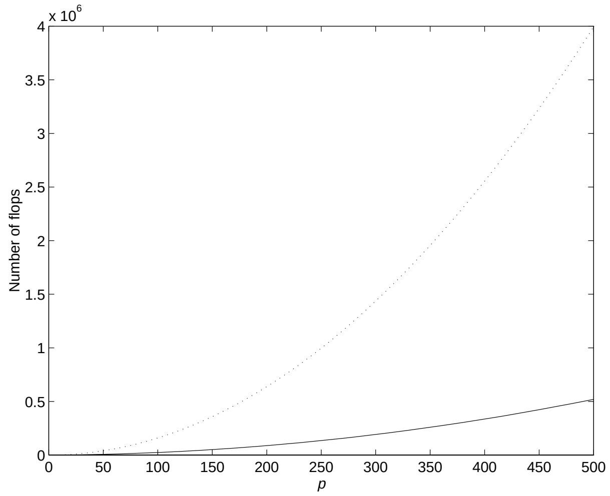

Figure 1: Number of flops versus p for $s = 5$. Levinson algorithm (dotted line) and OSR algorithm (solid line).

7. CONCLUSION

We have developed new algorithms for estimating multiple missing values of a multivariate stationary process. Our algorithms are not only order recursive but also step recursive. This twofold recursivity implies an important computational saving, as confirmed by the numerical simulations. Further complexity reduction can be achieved using highly parallel computational organizations which are particularly suitable for parallel processing. The structural properties of the proposed algorithms play an important role when a VLSI implementation is sought. Many alternative implementations of the algorithms presented in this paper are currently being investigated.

8. REFERENCES

[1] B. Abraham, "Missing observations in time series," *Commun. Statist. Theor. Meth. A,* vol. 10, no. 16, pp. 1643–1653, 1981.

[2] P. Bondon, D. P. Ruiz and A. Gallego, "Multiple-step linear prediction." Technical report, L2S/CNRS and Dept. FA/UGR, September 1997.

[3] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods,* New York: Springer Verlag, 1991.

[4] S. R. Brubacher and G. T. Wilson, "Interpolating time series with application to the estimation of holiday effects on electricity demand,"*Appl. Statist.* vol. 25, no. 2, pp. 107–116, 1976.

[5] J. Durbin, "The fitting of time series models," *Rev. Int. Stat.* vol. 28, pp. 233–244, 1960.

[6] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*. Englewood Cliffs, N. J. : Prentice-Hall, 1974.

[7] N. Levinson, "The Wiener RMS (root mean square) error criterion in filter design and prediction," *J. Math. Phys.,* vol. 25, no. 4, pp. 261–278, 1947.

[8] P. Whittle, "On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix," *Biometrika,* vol. 40, pp. 129–134, 1963.