NASA Technical Memorandum 100069

Linear Prediction of Stationary Vector Sequences

Yoram Baram

(NASA-TM-100069) LINEAR PREDICTION OF STATIONARY VECTOR SEQUENCES (NASA) 14 p CSCL 12B
N88-18366
Unclas
G3/66 0125852

March 1988

National Aeronautics and
Space Administration

NASA Technical Memorandum 100069

Linear Prediction of Stationary Vector Sequences

Yoram Baram, Ames Research Center, Moffett Field, California

March 1988

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

LINEAR PREDICTION OF STATIONARY VECTOR SEQUENCES

Yoram Baram*

SUMMARY

The class of all linear predictors of minimal order for a stationary vector-valued process is specified in terms of linear transformations on the associated Hankel covariance matrix. Two particular transformations, yielding computationally efficient construction schemes, are proposed.

1. INTRODUCTION

The prediction of a vector-valued stationary sequence possessing a linear representation of finite order can be approached, in principle, by first constructing such a representation from the process second order statistics and then matching a Kalman filter to it. The construction of a linear representation can be performed by spectral factorization, as suggested by Anderson [1] or by stochastic realization, as suggested by Faurre [2]. Given the covariance function of the process in the time domain, both approaches would be indirect, as the first requires the transformation to the frequency domain, while the second requires the intermediate solution of a different problem, namely, the deterministic input-output realization problem. The construction of the Kalman filter further requires the solution of a Riccati equation. More direct predictor constructions from the covariance function, suggested by Faurre [3] and by Son and Anderson [4], also require the solution of Riccati equations.

In this paper we propose a direct approach to the construction of linear predictors for stationary vector sequences from the covariance function. The approach, inspired by Akaike's coordinate-free realization concepts [5], is based on simple geometric principles. It suggests an explicit coordinate-dependent characterization of the class of all minimal order linear predictors, in terms of linear transformations on the Hankel covariance matrix associated with the sequence. It does not give rise to Riccati or Lyapunov equations. The selection of particular transformation matrices defines specific predictor construction techniques.

*Y. Baram is a Senior Research Associate of the National Research Council at the NASA Ames Research Center, Moffett Field, CA 94035, on sabbatical leave from the Department of Electrical Engineering, Technion, Israel Institute of Technology, Haifa 32000, Israel.

2. GENERALIZED FORMULATION

Let the covariance function of a zero-mean, full rank stationary sequence $y_n \in R^m$ be given by

$$R_k = E\{y_n y_{n-k}^T\} \tag{2.1}$$

where E denotes the expectation operation, and suppose that there exists a positive integer P and a set of scalars $a_1, \ldots, a_{P-1}$, $a_o = 1$, such that for any $n \geq P$

$$\sum_{k=0}^{P-1} a_k R_{n-k} = 0 \tag{2.2}$$

It is well known that when y_n is generated by a finite dimensional linear system driven by white noise, a relationship of the type (2.2) is satisfied.

The prediction problem is one of estimating the values of y_{n+k}, $k = 0,1,\ldots,$ given the values of y_{n-1}, $y_{n-2}, \ldots$. Let us denote by $Y_n^- = (y_{n-1}^T, y_{n-2}^T, \ldots, y_0^T)^T$ and $Y_n^+ = (y_n^T, y_{n+1}^T, \ldots)^T$ the past and the future vectors at time n. Let $y_{k,i}$ denote the ith element of y_k and let $y_{k|n-1}$ denote the linear mean-square projection of y_k on Y_n^-. Let us further denote by $(Y_n^+|Y_n^-)$ the space generated by $\{y_{k,i|n-1}, i = 1,\ldots,m, k \geq n\}$, by $Y_n^-(k) = (y_{n-1}^T, y_{n-2}^T, \ldots, y_{n-k}^T)^T$ and $Y_n^+(k) = (y_n^T, y_{n+1}^T, \ldots, y_{n+k-1}^T)^T$ the k-step past and future vectors at time n and by $Y_{n|n-1}(k)$ the linear mean-square projection of $Y_n^+(k)$ on Y_n^-. We have

$$Y_{n|n-1}^+(k) = R(k,n)[E\{Y_n^- Y_n^{-T}\}]^{-1} Y_n^-$$

where

$$R(k,n) = E\{Y_n^+(k) Y_n^-\} = \begin{bmatrix} R_1 & R_2 & \cdots & R_n \\ R_2 & R_3 & \cdots & R_{n+1} \\ \vdots & & & \\ R_k & R_{k+1} & \cdots & R_{n+k-1} \end{bmatrix}$$

It follows from (2.2) that for any $k \geq P$ the rows beyond the first $P \cdot m$ rows of $R(k,n)$ are linearly dependent on the previous ones. The space $(Y_n^+ | Y_n^-)$ is then spanned by

$$Y_{n|n-1}^+(P) = R(P,n)[E\{Y_n^- Y_n^{-T}\}]^{-1} Y_n^-$$

It can also be seen from (2.2) that the columns of $R(k,n)$ beyond the first $P \cdot m$ columns are linearly dependent on the previous ones. It follows that the first maximal set of linearly independent rows of $R(P,n)$ is indexed as the set of such rows of the matrix

$$R = E\{Y_n^+(P)[Y_n^-(P)]^T\} = \begin{bmatrix} R_1 & R_2 & \cdots & R_P \\ R_2 & R_3 & \cdots & R_{P+1} \\ \vdots & & & \\ R_P & R_{P+1} & \cdots & R_{2P-1} \end{bmatrix}$$

The subsequent analysis will require some further notation. Let A and B be any two matrices such that A has no fewer rows and no fewer columns than B. Let us denote by $r_B(A)$ the matrix whose rows are those rows of A, indexed as the first maximal set of linearly independent rows of B. Similarly, let us denote by $c_B(A)$ the matrix whose columns are those columns of A indexed as the first maximal set of linearly independent columns of B. We will find it convenient to write these row and column selection operations in terms of matrix products

$$r_B(A) \equiv r_B^A A$$

and

$$c_B(A) \equiv A c_B^{A^T}$$

where r_B^A and c_B^A are properly dimensioned, full rank matrices of zeros and ones. For instance, suppose that A has dimension 5×7 and that rows 1, 3, and 4 and columns 1, 2, and 5 of B form maximal independent sets. Then

$$r_B^A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

and

$$c_B^A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Since it may be assumed that the matrices involved have compatible dimensions, we will use the somewhat abbreviated notation

$$r_B \equiv r_B^A , \quad c_B \equiv c_B^A$$

Let M and N be arbitrary nonsingular square matrices of dimension $m \cdot P$ and let

$$X_n = MY^+_{n|n-1}(P)$$

and

$$Z_n = NY^-_n(P)$$

Let us further define

$$x_n = r_{MRN^T}X_n$$

and

$$z_n^T = Z_n^T c^T_{MRN^T}$$

It can be seen that x_n is a vector of minimal dimension which spans $(Y^+_n|Y^-_n)$ in the sense that each element of the latter space can be obtained by a linear combination of the elements of x_n. Clearly, x_{n+1} belongs to $(Y^+_{n+1}|Y^-_{n+1})$, which is included in $(Y^+_n|Y^-_{n+1})$. The latter may be decomposed as $(Y^+_n|Y^-_{n+1}) = (Y^+_n|Y^-_n) \otimes V_n$ where V_n is the space generated by

$$v_n = y_n - y_{n|n-1}$$

and $\otimes$ denotes the Cartesian product. Since V_n is orthogonal to $(Y_n^+|Y_n^-)$, it follows that there exist matrices A_n and B_n such that

$$x_{n+1} = A_n x_n + B_n v_n \tag{2.4}$$

Multiplying (2.4) on the right by z_n^T, taking expectation and noting that v_n is orthogonal to Y_n^-, hence, to z_n, we obtain

$$E\{x_{n+1} z_n^T\} = A_n E\{x_n z_n^T\} \tag{2.5}$$

where

$$\begin{aligned} E\{x_n z_n^T\} &= r_{MRN^T} M E\{Y_{n|n-1}^T (P)[Y_n^-(P)]^T\} N^T c_{MRN^T}^T \\ &= r_{MRN^T} MRN^T c_{MRN^T}^T \\ &= \bar{M}R\bar{N}^T \end{aligned} \tag{2.6}$$

where

$$\bar{M} = r_{MRN^T} M , \qquad \bar{N} = c_{MRN^T} N$$

It can be seen that $E\{x_n z_n^T\}$ is a non-singular matrix. We also have

$$\begin{aligned} E\{x_{n+1} z_n^T\} &= r_{MRN^T} M E\{Y_{n+1|n}(P)[Y_n^-(P)]^T\} N^T c_{MRN^T}^T \\ &= \bar{M}R^S\bar{N}^T \end{aligned} \tag{2.7}$$

where

$$R^S = E\{Y_{n+1}^+(P)[Y_n^-(P)]^T\} = \begin{bmatrix} R_2 & R_3 & \cdots & R_{P+1} \\ R_3 & R_4 & \cdots & R_{P+2} \\ \vdots & & & \\ R_{P+1} & P_{P+2} & \cdots & R_{2P} \end{bmatrix} \tag{2.8}$$

Substituting (2.6) and (2.7) into (2.5), we obtain

$$A_n = A = \bar{M}R^S\bar{N}^T(\bar{M}R\bar{N}^T)^{-1}$$

It follows from the definition of x_n that there exists a matrix C_n such that

$$y_n = C_n x_n + v_n \tag{2.9}$$

Multiplying (2.9) on the right by z_n^T and taking expectation, we obtain

$$C_n = E\{y_n z_n^T\}[E\{x_n z_n^T\}]^{-1} \tag{2.10}$$

where

$$E\{y_n z_n^T\} = (MRN^T)_1 c^T_{MRN^T}$$

where $(MRN^T)_1$ is the first block row of MRN^T. Hence,

$$C_n = C = (MRN^T)_1 c^T_{MRN^T}(\bar{M}R\bar{N}^T)^{-1}$$

Multiplying (2.4) on the right by v_n^T and taking expectation, noting that v_n is orthogonal to x_n, we obtain

$$B_n = B = E\{x_{n+1}v_n^T\}[E\{v_n v_n^T\}]^{-1} \tag{2.11}$$

where

$$E\{x_{n+1}v_n^T\} = E\{x_{n+1}y_n^T\} - E\{x_{n+1}x_n^T\}C^T$$

$$= \bar{M}\underline{R} - A\Pi C^T \tag{2.12}$$

and

$$E\{v_n v_n^T\} = E\{(y_n - Cx_n)y_n^T\}$$
$$= R_o - C\Pi C^T \tag{2.13}$$

where

$$\Pi = E\{x_n x_n^T\}$$

and

$$\underline{R} = (R_1^T \ R_2^T \cdots R_P^T)^T \tag{2.14}$$

We next derive the term ΠC^T, which appears in both (2.12) and (2.13). Noting that

$$R_k = E\{y_{n+k} y_n^T\} = CA^k \Pi C^T$$

it can be seen that

$$\Theta \Pi C^T = \underline{R}$$

Where $\underline{R}$ is defined by (2.14) and

$$\Theta = \begin{bmatrix} CA \\ CA^2 \\ \vdots \\ CA^P \end{bmatrix} \tag{2.15}$$

Since $Y_{n|n-1}(P) = \Theta x_n$ and since the covariance rank of x_n is the same as that of $Y_{n|n-1}(P)$, it follows that Θ is full rank. Hence,

$$\Pi C^T = (\Theta^T \Theta)^{-1} \Theta^T \underline{R}$$

which, substituted in (2.12) and (2.13), completes the derivation of the matrix B, defined by (2.11). In summary, the generalized predictor is given by

$$\begin{aligned} x_{n+1} &= Ax_n + Bv_n \,, \qquad x_o = 0 \\ y_{n+k|n-1} &= CA^k x_n \end{aligned} \tag{2.16}$$

where

$$A = \bar{M}R^{s}\bar{N}^{T}(\bar{M}R\bar{N}^{T})^{-1} \tag{2.17}$$

$$C = (MRN^{T})_{1}c^{T}_{MRN^{T}}(\bar{M}R\bar{N}^{T})^{-1} \tag{2.18}$$

$$B = [\bar{M} - A(\theta^{T}\theta)^{-1}\theta^{T}]\underline{R}[R_{o} - C(\theta^{T}\theta)^{-1}\theta^{T}\underline{R}] \tag{2.19}$$

and

$$v_n = y_n - Cx_n$$

The class of all minimal order linear predictors for the sequence y_n is now defined by (2.16). A particular member of the class is specified by selection of the nonsingular matrices M and N. Two particular selections are suggested in the following section. We note that

$$\begin{aligned} x_{n+1} &= Ax_n + Bv_n \\ y_n &= Cx_n + v_n \end{aligned} \tag{2.20}$$

With A, B, C, and v_n as defined above, is a minimal realization of the sequence y_n.

3. SPECIFIC CONSTRUCTIONS

A. Original Coordinates

When the matrices M and N are taken to be identity matrices, the vectors $Y_n^-(P)$ and $Y_n^+(P)$ are left in their original coordinates. The minimal predictor for this choice is specified in the preceding section, with MRN^T and MR^sN^T replaced simply by R and R^s. The predictor is essentially defined by maximal sets of linearly independent rows and linearly independent columns of R. The following procedure selects a maximal set of such rows. Extension to the selection of linearly independent columns is immediate.

The procedure consists of checking the rank deficiency of matrices R_i, consisting of the first i-1 linearly independent rows of R and the row following them. If R_i is rank-deficient, its last row is replaced by the next row of R and the rank check is repeated. If R_i has full-column rank, the next row of R is attached to R_i to form R_{i+1} and a rank check is performed. This procedure is continued until rank $R_i = M$ or until all the rows of R have been checked. The final R_i will contain a maximal set of linearly independent rows of R. This procedure can be extended to the construction of a linear predictor or a minimal realization from a sample covariance sequence, employing a recently proposed statistical rank test method [6].

B. Canonical Coordinates

Construction in the original coordinates requires a sequence of rank checks on submatrices of the Hankel covariance matrix R. A computationally more attractive technique may be obtained by selecting the matrices M and N so as to diagonalize the matrix MRN^T. Let M and N be the matrices of normalized eigenvectors of RR^T and R^TR, respectively, corresponding to the nonzero eigenvalues. These may be obtained directly from the singular value decomposition of R (see, e.g., [7]). It can be readily seen that

$$r_{MRN^T} = c_{MRN^T} = [I_p \quad 0] \tag{3.1}$$

Denoting

$$S = \begin{bmatrix} s_1 & & & 0 \\ & s_2 & & \\ & & \ddots & \\ 0 & & & s_p \end{bmatrix}$$

where $s_1,\ldots,s_p$ are the squared nonzero singular values of R, the predictor is now specified by the matrices

$$A = [I_p \quad 0]MR^sN^T \begin{bmatrix} I_p \\ 0 \end{bmatrix} S^{-1} \tag{3.2}$$

$$C = [s_1 \quad 0 \cdots 0]S^{-1} \tag{3.3}$$

and B is defined by (2.19), with r_{MRN^T} defined by (3.1), Θ defined by (2.15), and A and C defined by (3.2) and (3.3). We note that while the above procedure is computationally more effective than the one performed in the original coordinates, when the exact covariance function is known, it is not presently extendable by means of statistical inference to the case where only a sample covariance sequence is given. We also note that the above construction is different from the canonical variates realization method suggested by Larimore [8], which produces an approximate representation for the process, even when the latter possesses a finite order realization. The present construction produces an exact realization of the form (2.20) with the parameters specified above.

4. CONCLUSION

A direct approach to the construction of a minimal order linear predictor for a stationary vector sequence from its covariance function has been proposed. The class of all such predictors has been characterized in terms of linear transformations on the Hankel covariance matrix. Two specific construction schemes have been specified.

REFERENCES

1. B. D. O. Anderson, "The Inverse Problem of Stationary Covariance Generation," J. Stat. Phys., Vol. 1, No. 1, pp. 133-147, 1969.

2. P. L. Faurre, "Identification par minimisation d'une representation markovienne de processus aleatoire," Symp. on Optimization, Nice, June 1969, Springer-Verlag Lecture Notes on Mathematics, 132.

3. P. L. Faurre, "Stochastic Realization Algorithms," in System Identification, R. K. Mehra and D. G. Lainiotis eds., Academic Press, New York, 1976.

4. L. H. Son and B. D. O. Anderson, "Design of Kalman Filters Using Signal Model Output Statistics," Proc. IEE, Vol. 120, No. 2, pp. 312-318, February 1973.

5. H. Akaike, "Stochastic Theory of Minimal Realization," IEEE Trans. on Automatic Control, Vol. AC-19, No. 6, December 6, 1974.

6. Y. Baram and B. Porat, "Identification of Minimal Order State Space Models from Stochastic Input-Output Data," to appear in SIAM J. on Contr. Optimization, Vol. 26, No. 1, January 1988.

7. G. Golub and W. Kahan, "Calculating the Singular Values and Pseudo-Inverse of a Matrix," SIAM J. Numerical Analysis, Ser. B, Vol. 2, pp. 205-224, 1965.

8. W. E. Larimore, "System Identification, Reduced Order Filtering and Modelling via Canonical Variates Analysis," Proc. Amer. Contr. Conf., H. S. Rao and T. Dorats, eds., IEEE, New York, pp. 445-451, 1983.

NASA
National Aeronautics and Space Administration

Report Documentation Page

1. Report No. NASA TM-100069	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Linear Prediction of Stationary Vector Sequences		5. Report Date March 1988
		6. Performing Organization Code
7. Author(s) Yoram Baram		8. Performing Organization Report No. A-88087
		10. Work Unit No. 505-66-11
9. Performing Organization Name and Address Ames Research Center Moffett Field, CA 94035		11. Contract or Grant No.
		13. Type of Report and Period Covered Technical Memorandum
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, DC 20546		14. Sponsoring Agency Code

15. Supplementary Notes

Point of Contact: Heinz Erzberger, MS 210-9, Ames Research Center, Moffett Field, CA 94035, (415) 694-5425 or FTS 464-5425

16. Abstract

The class of all linear predictors of minimal order for a stationary vector-valued process is specified in terms of linear transformations on the associated Hankel covariance matrix. Two particular transformations, yielding computationally efficient construction schemes, are proposed.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
System compensation Inversion identification	Unclassified - Unlimited Subject category - 66

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	12	A02