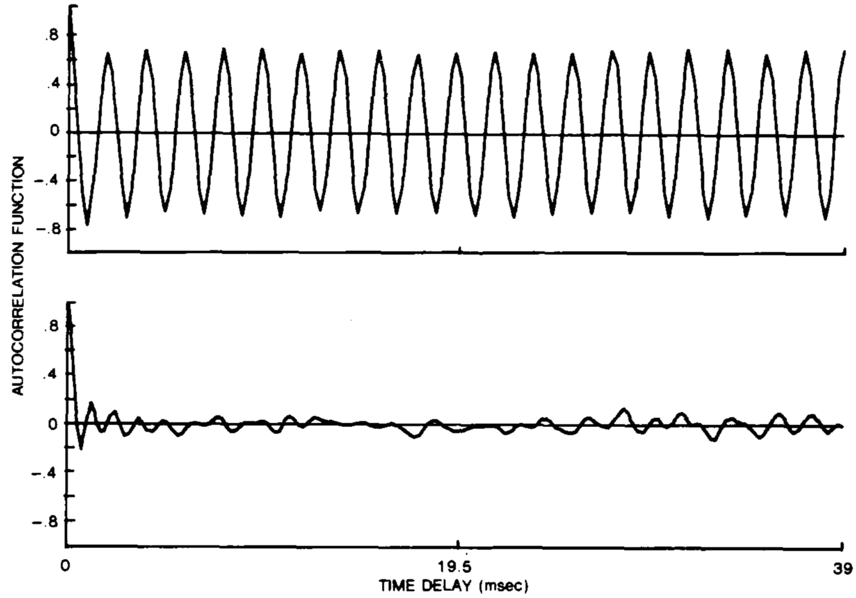

Fig. 4. Autocorrelation function: top, broad-band signal and sinusoid; bottom, autocorrelation function of broad-band signal after sinusoid has been removed by a nonlinearity.

tionship was defined to measure the effectiveness of the INL concept. Two examples were given to demonstrate the method.

REFERENCES

[1] R. Dwyer, "A technique for improving detection and estimation of signals contaminated by under ice noise," *J. Acoust. Soc. Amer.*, vol. 74, no. 1, July 1983.

[2] F. R. DiNapoli *et al.*, "Tristen/FRAM II Cruise Report East Arctic, April 1980," NUSC Doc. TD 6457, Apr. 13, 1981.

[3] R. F. Dwyer, "Detection of non-Gaussian signals by frequency domain Kurtosis estimation," in *1983 ICASSP Conv. Rec.*

[4] G. C. Carter, A. H. Nuttall, and P. G. Cable, "The smoothed coherent transform," *Proc. IEEE* (Lett.), vol. 61, pp. 1497–1498, Oct. 1973.

[5] J. Hassab and R. Boucher, "Performance of the generalized cross correlator in the presence of a strong spectral peak in the signal," *IEEE Trans. Accoust., Speech, Signal Process.*, vol. ASSP-29, no. 3, June 1981.

[6] D. Tufts, R. Kumaresan, and I. Kirsteins, "Data adaptive signal estimation by singular value decomposition of a data matrix," *Proc. IEEE*, vol. 70, no. 6, June 1982.

On the Finite Maximum Entropy Extrapolation

MEIR FEDER AND EHUD WEINSTEIN

It is shown that an autoregressive (AR) extrapolation of a given set of correlation lags by any finite number maximizes the entropy (i.e., the determinant of the correlation matrix) of the corresponding segment of the time series.

Manuscript received October 3, 1983; revised January 16, 1984. It is a Woods Hole Oceanographic Institution Contribution 5502.

M. Feder is with the Department of Ocean Engineering, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA, on leave from the Department of Electronic Systems, Faculty of Engineering, Tel-Aviv University, Ramat-Aviv, Israel.

E. Weinstein is with the Department of Ocean Engineering, Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA, and with the Department of Electronic Systems, Faculty of Engineering, Tel-Aviv University, Ramat-Aviv, Israel.

In many signal processing problems such as filtering, system identification, and spectral analysis it is desired to extrapolate a measured set of correlation lags of the observed time series so as to increase the resolution capabilities. One possible approach is to use the maximum entropy method (MEM). According to this criterion, one extrapolates the limited available correlation data into the unknown region of the time domain so that the entropy of the underlying random process is maximized. In this way one ensures that the fewest assumptions have been made concerning the unmeasured data.

The problem we are addressing in this letter may be stated as follows: Given partial knowledge of the first N correlation lags, find an extrapolation by some arbitrary number M using the maximum entropy principle. Assuming that the correlation sequence corresponds to a weakly stationary discrete-time Gaussian random process, the entropy of the corresponding $(N + M)$ segment of the time series is given by [1]

$$H = \frac{N+M}{2} \log(2\pi e) + \frac{1}{2} \log(\det R^{N+M}) \tag{1}$$

where $\det R^{N+M}$ is the determinant of R^{N+M}. R^{N+M} is the $(N + M) \times (N + M)$ matrix whose (i, j) element is given by

$$R_{ij}^{N+M} = r_{i-j}, \qquad i, j = 0, 1, \cdots, (N + M - 1) \tag{2}$$

and r_ℓ is the correlation value at the ℓth time lag.

Since the first term in (1) is a constant and the logarithmic function is monotonic, the maximum entropy problem can be stated as follows: Given the first N correlation lags $(r_0, r_1, \cdots, r_{N-1})$, find an extension $(r_N, r_{N+1}, \cdots, r_{N+M-1})$ such that

$$\underset{(r_N, r_{N+1}, \cdots, r_{N+M-1})}{\mathrm{Max}} [\det R^{N+M}]. \tag{3}$$

In [2] it has been shown that the solution to the one-dimensional problem (i.e., $M = 1$)

$$\underset{r_N}{\mathrm{Max}} [\det R^{N+1}] \tag{4}$$

0018-9219/84/1100-1660$01.00 ©1984 IEEE

is given by

$$r_N = \sum_{i=1}^{N-1} a_i^N r_{N-i} \tag{5}$$

where

$$\{a_i^N\}_{i=1}^{N-1}$$

are the coefficients of the prediction filter of order N obtained from

$$\begin{bmatrix} r_0 & r_1 & \cdots & r_{(N-2)} \\ r_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & r_1 \\ r_{(N-2)} & \cdots & r_1 & r_0 \end{bmatrix} \begin{bmatrix} a_1^N \\ a_2^N \\ \vdots \\ a_{N-1}^N \end{bmatrix} = - \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_{N-1} \end{bmatrix} \tag{6}$$

Using the AR extrapolation in (5) recursively to generate higher order correlation lags corresponds to successive one-step maximizations (i.e., one first maximizes $\det R^{N+1}$ with respect to r_N, then substitutes the resulting r_N into R^{N+2} and maximizes $\det R^{N+2}$ with respect to r_{N+1}, etc.). In general, maximizing the entropy at each step is different from maximizing $\det R^{N+M}$ with respect to all M variables at once. However, in this case, both procedures are equivalent as stated by the following theorem.

Theorem:

The solution to (3) is obtained using the extrapolation formula

$$r_k = \sum_{i=1}^{N-1} a_i^N r_{k-i}, \qquad k = N, N+1, \cdots, (N+M-1) \tag{7}$$

for any given $M \geqslant 1$.

Proof: We first observe that

$$\frac{\det R^{N+M}}{\det R^N} = \prod_{k=1}^{M} \frac{\det R^{N+k}}{\det R^{N+k-1}} = \prod_{k=1}^{M} P_{N+k}. \tag{8}$$

Since $\det R^N$ depends only on the first N correlation lags, maximizing $\det R^{N+M} / \det R^N$ with respect to $r_N, r_{N+1}, \cdots, r_{N+M-1}$ will yield the desired result. In this setting it is sufficient to find an extrapolation to the M unknown correlation lags which simultaneously maximizes all the terms P_{N+k} in the product, i.e., it is sufficient to solve

$$\max_{(r_N, r_{N+1}, \cdots, r_{N+k-1})} P_{N+k}, \qquad k = 1, 2, \cdots, M. \tag{9}$$

In [3] it has been shown that P_{N+k} is the minimum attainable mean square error (m.s.e.) in the prediction filter of order $(N+k)$. Obviously, $P_{N+k} \leqslant P_N$ for all positive k and any combination of $r_N, r_{N+1}, \cdots, r_{N+k-1}$—otherwise the prediction filter of order N could be used to achieve a smaller m.s.e.[1] If, however, the M unknown correlation lags are generated using (7), the indicated upper bound is reached simultaneously for all k, i.e., $P_{N+k} = P_N$, $k \geqslant 1$. This happens because in that case the prediction filter of order $(N+k)$ coincides with the prediction filter of order N, i.e.,

$$a_i^{N+k} = \begin{cases} a_i^N, & i = 1, 2, \cdots, (N-1) \\ 0, & i \geqslant N. \end{cases} \tag{10^2}$$

The AR extrapolation (7), therefore, is the solution to (9) and the theorem is proved.

[1]This can also be observed from the Levinson's recursion formulas for solving the Yule–Walker equations.

[2]Equation (10) is immediately obtained by writing (6) for the $(N+k)$ case and exploiting the AR relation of (7).

Discussion

In a more general context [4], it has been shown that the inverse of the extended matrix, R^{N+M} in our case, that maximizes the determinant must assume the general form

$$(R^{N+M})^{-1} = \left[\begin{array}{cccc} & & \overbrace{\begin{matrix} 0 & 0 & \cdots & 0 \\ & 0 & \ddots & \vdots \\ & & \ddots & \vdots \\ & & & 0 \end{matrix}}^{M} \\ \underbrace{\begin{matrix} 0 & & & \\ 0 & 0 & & \\ \vdots & & \ddots & \\ 0 & \cdots & \cdots & 0 \end{matrix}}_{M} & & \end{array}\right] \tag{11}$$

Equation (11) reads that $(R^{N+M})_{ij}^{-1} = 0$ for $|i-j| \geqslant N$. A matrix of this general form is called a band matrix. To show that the above theorem is consistent with that result, we first observe that R^{N+M} can always be diagonalized by [5]

$$A^T R^{N+M} A = \operatorname{diag}(P_{N+M-1}, P_{N+M-2}, \cdots, P_0) \tag{12}$$

where A^T is the upper-triangular matrix

$$A^T = \begin{bmatrix} 1 & a_1^{N+M} & a_2^{N+M} & \cdots & a_{N+M-1}^{N+M} \\ 0 & 1 & a_1^{N+M-1} & \cdots & a_{N+M-2}^{N+M-1} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & a_1^2 \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix}. \tag{13}$$

It follows that

$$R^{N+M} = (A^T)^{-1} \Lambda A^{-1}. \tag{14}$$

Hence

$$(R^{N+M})^{-1} = A \Lambda^{-1} A^T. \tag{15}$$

If R^{N+M} is extended using the AR extrapolation formula (7), then the second line in (10) reads that $a_i^{N+k} = 0$ whenever $i \geqslant N$. In that case, (13) assumes the form

$$A^T = \left[\begin{array}{ccc} 1 & & \overbrace{\begin{matrix} 0 & \cdots & 0 \\ & \ddots & \vdots \\ & & 0 \end{matrix}}^{M} \\ & 1 \ddots & \\ (0) & & \ddots \; 1 \end{array}\right] \tag{16}$$

where the notation (0) indicates that all the elements below the main diagonal are zero. Substituting (16) into (15) one immediately obtains

$$
(R^{N+M})^{-1} = \begin{bmatrix} 1 & & & & (0) \\ & \ddots & & & \\ 0 & & \ddots & & \\ \vdots & \ddots & & \ddots & \\ \underbrace{0 \cdots 0}_{M} & & & & 1 \end{bmatrix} \begin{bmatrix} 1/P_{N+M-1} & & (0) \\ & \ddots & \\ (0) & & 1/P_0 \end{bmatrix} \begin{bmatrix} 1 & & & & \overbrace{0 \cdots 0}^{M} \\ & \ddots & & \ddots & \vdots \\ & & \ddots & & 0 \\ & & & \ddots & \\ (0) & & & & 1 \end{bmatrix}
$$

$$
= \begin{bmatrix} & & & & \overbrace{0 \cdots 0}^{M} \\ & & & \ddots & \vdots \\ & & & & 0 \\ 0 & & & & \\ \vdots & \ddots & & & \\ \underbrace{0 \cdots 0}_{M} & & & & \end{bmatrix} \tag{17}
$$

REFERENCES

[1] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication.* Urbana, IL: Univ. of Illinois Press, 1959.

[2] A. Van Den Bos, "Alternative interpretation of maximum entropy spectral analysis," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 493–494, July 1971.

[3] K. Yao, "On the direct calculations of MMSE of linear realizable estimator by Toeplitz form method," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 95–97, Jan. 1971.

[4] H. Dym and I. Gohberg, "Extensions of band-matrices with band-inverses," *Linear Algebra Appl.*, vol. 36, pp. 1–24, 1981.

[5] J. P. Burg, "Maximum entropy spectral analysis," Ph.D. dissertation, Stanford University, Stanford, CA, 1975.

Book Reviews

The following reviews were selected from those recently published in various IEEE TRANSACTIONS *and Group/Society Magazines and Newsletters. They are reprinted here to make them conveniently available to the many readers who otherwise might not have ready access to them. Each review is followed by an identification of its original source.*

Optische Nachrichtentechnik—G. Grau. (Berlin, Germany: Springer-Verlag, 1981, 144 pp.) *Reviewed by H. Melhior, Swiss Federal Institute of Technology, Zürich, Switzerland.*

This book, titled *Optical Communication*, is written for electrical engineers and engineering students and gives an introduction into the field of fiber optical communication. The book deals with optical fibers, semiconductor light sources, photodetectors, and optical communication links; it represents the status of this fast-moving field around 1980.

Starting with the material properties of silica, the wave propagation, attenuation, and dispersion of monomode and multimode fibers are treated in great mathematical detail, followed by a few brief remarks about the fabrication of fibers. The chapters on light sources describe light-emission diodes and semiconductor lasers and cover typical construction types and a number of their electro-optical characteristics. Photodiodes and avalanche photodiodes are presented, as well as noise and sensitivity calculations of entire receivers for analog and digital signals. Fiber splices and couplers are mentioned. A brief chapter on applications concludes the book.

This book is undoubtedly one of the best introductions to the field of fiber optical communications available in the German language. It emphasizes mathematical calculations. In this reviewer's opinion, it takes some effort by the reader to work out the physical principles underlying the various mathematical treatments and to extract the data needed for practical applications. Nonetheless, it is a careful introduction for students as well as developers of fiber-optical communication systems.

Reprinted from *IEEE Circuits and Systems Magazine*, vol. 5, no. 2, p. 13, June 1983.

Multivariable System Theory and Design—R. V. Patel and N. Munro. (Oxford, England: Pergamon, 1982, 374 pp.) *Reviewed by George C. Verghese, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139.*

This book deals with methods of analysis and control design for multi-input-output, *linear, time-invariant* dynamical systems. The area continues to be an active and important one in research and applications. The authors' primary intent, as stated in the preface, is to present a comprehensive and up-to-date treatment of a selection of recent developments in this area, in a textbook suitable for graduate control engineering students. As prerequisites for studying the book, they list an introductory course in the state-space approach to dynamical systems, basic undergraduate courses in linear algebra and complex variables, and preferably a course in classical control theory as well (though the latter is claimed not to be necessary for understanding the material in the book); the mathematical level of the book is intentionally kept fairly basic. It is also intended to be useful for research and reference purposes.

The following listing of chapter headings and lengths will convey some idea of the contents and level of this book. Following an introductory chapter (chap. 1, 18 pp.) that consists of a brief historical perspective on the area (with 161 references), the book contains chapters titled: "Multivariable System Representations" (chap. 2, 18 pp.); "Controllability, Observability, and Canonical Forms" (chap. 3, 34 pp.); "Poles and Zeros of Multivariable Systems"