The Barankin bound was derived from the Schwarz inequality, in a manner much like (6), by Kiefer [7]. One of the reviewers pointed out that the Barankin formulation can be specialized to include all the bounds for unbiased estimation, e.g., the Bhattacharyya bounds. The formulation in this correspondence is limited since only first derivatives of H are considered. On the other hand, the formulation of (6) is not limited to unbiased estimation.

Remarks

It can be shown that Zacks' bound (15) is smaller (looser) than the C–R bound (18), unless the vector V happens to be an eigenvector of D^T. In that case the bounds are equal. This occurs for all V if $D = I$.

In signal parameter estimation problems one is almost always forced to assume $D = I$, an identity matrix. Otherwise, information on the derivatives in D must be available.

ACKNOWLEDGMENT

The connections to the Barankin bound were pointed out by a referee.

REFERENCES

[1] L. P. Seidman, "Performance limitations and error calculations for parameter estimation," *Proc. IEEE*, vol. 58, pp. 644–652, May 1970.
[2] C. R. Rao, "Minimum variance and the estimation of several parameters," *Proc. Cambridge Phil. Soc.*, vol. 43, pp. 280–283, 1946.
[3] S. Zacks, *The Theory of Statistical Inference*. New York: Wiley, 1971, pp. 194–200.
[4] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*. New York: Wiley, 1968.
[5] E. W. Barankin, "Locally best unbiased estimators," *Ann. Math. Statist.*, vol. 20, pp. 477–501, 1949.
[6] R. J. McAulay and E. M. Hofstetter, "Barankin bounds on parameter estimation," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 669–676, Nov. 1971.
[7] J. Kiefer, "On minimum variance estimators," *Ann. Math. Statist.*, vol. 23, pp. 627–629, 1952.

Cepstrum Discrimination Function

R. G. SMITH

Abstract—**An expression is derived for the function that governs the discrimination by the power cepstrum against components at large time delays. The function has been found to be useful in normalizing cepstrum displays.**

Cepstrum analysis is an appropriate technique for investigation of the different arrival times of continuous signals received in a multipath environment. A problem does arise, however, in the display of the results due to the use of data records of finite length in the intermediate spectrum computation. Components at large time delay are attenuated relative to those at short time delay [1] in a manner similar to that encountered in estimation of correlation functions [2]. In this correspondence an expression is derived for the exact form of the discrimination function for the cepstrum of a process consisting of a noiselike signal received as a primary component accompanied by a delayed, possibly attenuated, secondary component. The discrimination of the cepstrum at large time delay is found to be more severe than in the corresponding case of autocorrelation. A knowledge of the functional form of the discrimination is useful in that it allows simple rescaling of the computed cepstral values prior to display.

Manuscript received June 10, 1974; revised December 11, 1974.
The author is with Defense Research Establishment Atlantic, Dartmouth, N. S., Canada.

This equalizes the contributions made by arrivals at all time delays and prevents domination of the display by arrivals at small time delay.

We consider a signal in the discrete form (1) that arises, for example, in propagation of continuous underwater acoustic signals via both a direct and a surface reflected path

$$s_k = x_k + \alpha x_{k-m}, \qquad k = 0, 1, \cdots, N-1 \tag{1}$$

where x_k is the signal received via the direct path, αx_{k-m} is the signal received via the surface reflected path with a relative time delay m and attenuation coefficient α, and s_k is the composite received signal, neglecting the effects of noise. Furthermore, x_k is assumed to be a stationary band-limited process with variance σ_x^2 and uniform spectral content, such that successive samples can be assumed to be uncorrelated.

In order to compute the cepstrum discrimination function, we will first compute the autocorrelation function of the signal described by (1) and then the power cepstrum. In this way we can compare the discrimination by the two functions against components at large time delay.

We will use the same definition of the power cepstrum as used in [1]; that is, the power cepstrum of a signal is the power spectrum of the logarithm of the power spectrum of the signal.

The discrete Fourier transform (DFT) of the signal is first computed using appropriate normalization

$$F_r = N^{-1/2} \sum_{k=0}^{N-1} (x_k + \alpha x_{k-m}) \exp\left(\frac{-j2\pi kr}{M}\right), \qquad r = 0, 1, \cdots, M-1 \tag{2}$$

where the N point data sequence has been augmented by $(M - N)$ zeros (generally $M \geq 2N$) to avoid aliasing effects.

Equation (2) can be rewritten as follows

$$\begin{aligned} F_r = N^{-1/2} \Bigg[&\sum_{k=N-m}^{N-1} x_k \exp\left(\frac{-j2\pi kr}{M}\right) \\ &+ \sum_{k=0}^{N-m-1} x_k \left(1 + \alpha \exp\left(\frac{-j2\pi mr}{M}\right)\right) \exp\left(\frac{-j2\pi kr}{M}\right) \\ &+ \alpha \exp\left(\frac{-j2\pi mr}{M}\right) \sum_{k=-m}^{-1} x_k \exp\left(\frac{-j2\pi kr}{M}\right) \Bigg], \\ &r = 0, 1, \cdots, M-1. \end{aligned} \tag{3}$$

An estimate of the spectrum S_r of s_k is computed as follows

$$\hat{S}_r = \overline{F_r^* F_r} \tag{4}$$

where $\overline{\qquad}$ denotes time averaging over many records. It is assumed that with sufficient averaging $\hat{S}_r$ will approach the ensemble average S_r, where

$$S_r = \sigma_x^2(1 + \alpha^2) + 2\alpha \left(\frac{N-m}{N}\right) \sigma_x^2 \cos\left(\frac{2\pi mr}{M}\right), \qquad r = 0, 1, \cdots, M-1. \tag{5}$$

Now, an estimate of the discrete autocorrelation function is obtained simply by inverse Fourier transformation of the estimated spectrum. It is apparent that there will be two terms in this estimate; one centered at zero time delay due to the constant term, and the other centered at time delay m due to the cosine term that is associated with the second arrival. It is also apparent that the amplitude of the second term decreases linearly with time delay. This discrimination is a familiar result of the use of data records of finite length.

It is now relatively simple to derive the discrimination function for the cepstrum. The logarithm of the spectrum must be considered first. Hence (5) may be rewritten (with some manipulation) as follows

$$\log(S_r) = \log\left\{1 + \left(\frac{2\alpha}{1+\alpha^2}\right)\left(\frac{N-m}{N}\right)\cos\left(\frac{2\pi mr}{M}\right)\right\} + \log(\sigma_x^2(1+\alpha^2)), \qquad r = 0, 1, \cdots, M-1. \tag{6}$$

Only the first term is of interest in considering the discrimination function. We will expand this term in a power series of the form (7) and simplify with the aid of (8) and (9) [3] to make its effect more readily apparent

$$\log(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}x^n}{n}, \qquad -1 < x < 1 \tag{7}$$

$$\cos^{2k}\theta = 2^{1-2k}\left[\sum_{l=0}^{k-1}\binom{2k}{l}\cos[2(k-l)\theta] + \frac{1}{2}\binom{2k}{k}\right] \tag{8}$$

$$\cos^{2k+1}\theta = 2^{-2k}\sum_{l=0}^{k}\binom{2k+1}{l}\cos[(1+2(k-l))\theta]. \tag{9}$$

Use of (7)–(9) on the first term of (6) leads to the following equation, where terms of the series have been rearranged and Q_r represents the first term of (6)

$$Q_r = \sum_{\substack{n=1\\ \text{odd}}}^{\infty} n^{-1}2^{1-n}\beta^n\left[\sum_{\substack{p=1\\ \text{odd}}}^{n}\binom{n}{\frac{n-p}{2}}\cos\left(\frac{2\pi pmr}{M}\right)\right] - \sum_{\substack{n=2\\ \text{even}}}^{\infty} n^{-1}2^{1-n}\beta^n\left[\sum_{\substack{p=2\\ \text{even}}}^{n}\binom{n}{\frac{n-p}{2}}\cos\left(\frac{2\pi pmr}{M}\right)\right] - \frac{1}{2}\sum_{\substack{n=2\\ \text{even}}}^{\infty} n^{-1}2^{1-n}\beta^n\binom{n}{\frac{n}{2}}, \qquad r = 0, 1, \cdots, M-1 \tag{10}$$

where

$$\beta = \left(\frac{2\alpha}{1+\alpha^2}\right)\left(\frac{N-m}{N}\right). \tag{11}$$

The form of (10) demonstrates that there will be terms in the power cepstrum associated with multiples of the time delay m, that is, terms centered at $m, 2m, 3m, \cdots$. The fundamental term (the term associated with delay m) is of primary importance and summing all terms of the form $\cos(2\pi mr/M)$ yields

$$Q_{r|_m} = \sum_{\substack{n=1\\ \text{odd}}}^{\infty} n^{-1}2^{1-n}\beta^n\binom{n}{\frac{n-1}{2}}\cos\left(\frac{2\pi mr}{M}\right), \qquad r = 0, 1, \cdots, M-1. \tag{12}$$

Hence the coefficient of the term associated with time delay m in the power cepstrum is the following (where the magnitude square arises due to the definition of the power cepstrum)

$$C_m = \left|\sum_{\substack{n=1\\ \text{odd}}}^{\infty} n^{-1}2^{1-n}\beta^n\binom{n}{\frac{n-1}{2}}\right|^2. \tag{13}$$

The normalized form C_m/C_o is plotted in Fig. 1 as a function of the normalized time delay m/N for various values of α and $N = 256$.

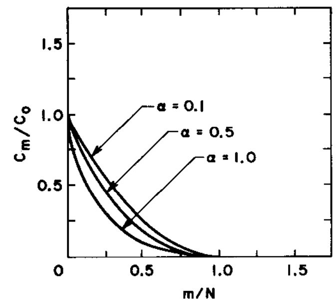

Fig. 1. Cepstrum discrimination function, $N = 256$.

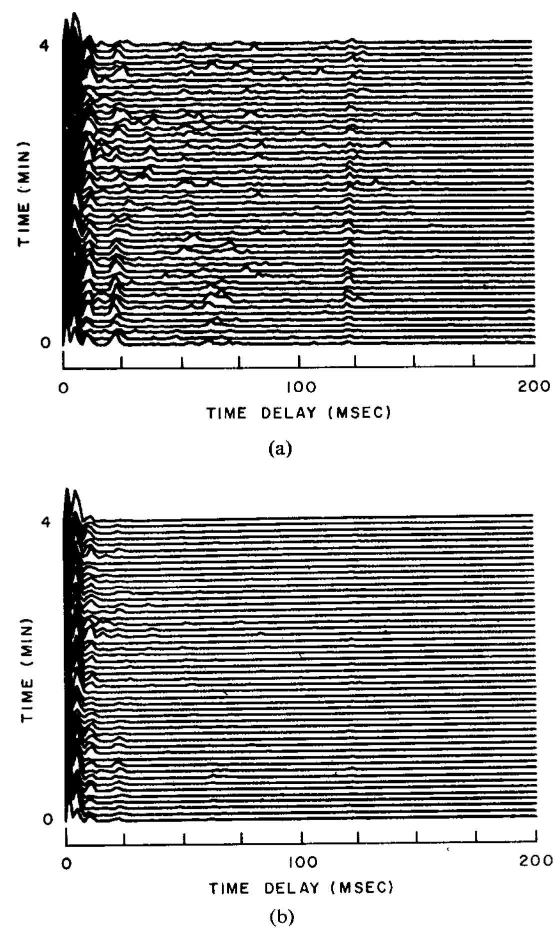

Fig. 2. Sample power cepstrum displays: (a) rescaled to remove discrimination; (b) normal.

The expression for C_m in (13) allows simple rescaling of the output of a power cepstrum processor prior to display, thus avoiding complete domination of the display by components at small time delay. This is demonstrated in the sample displays of Fig. 2, which show a time history of the power cepstrum of an underwater acoustic signal propagated via multiple paths.

Each succeeding time interval is indicated by a horizontal trace and cepstrum magnitude by linear deflection modulation. In part (a) of the figure, the display has been rescaled to remove discrimination according to the $\alpha = 1$ curve of Fig. 1, and the component of interest, at a delay of approximately 120 ms, is evident. In part (b), no rescaling has been done and the component is barely visible.

Equation (13) has been derived on the basis of a signal with uniform spectral content. Signals not meeting this requirement may lead to a more severe discrimination in that deviation from nonuniformity of the spectrum is generally interpreted by the power cepstrum as further small time delay content, thus adding to the problem. In such cases, the discrimination function should be used together with appropriate prewhitening techniques, such as liftering [4].

References

[1] R. C. Kemerait and D. G. Childers, "Signal detection and extraction by cepstrum techniques," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 745–759, Nov. 1972.
[2] J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures*. New York: Wiley, 1971, p. 313.
[3] V. Mangulis, *Handbook of Series for Scientists and Engineers*. New York: Academic, 1965, p. 19.
[4] B. P. Bogert, M. J. R. Healy, and J. W. Tukey, "The quefrency alanysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe-cracking," in *Proceedings of the Symposium on Time Series Analysis*, M. Rosenblatt, Ed. New York: Wiley, 1963, pp. 209–243.

A Note on the Use of Chandrasekhar Equations for the Calculation of the Kalman Gain Matrix

ROBERT F. BRAMMER, MEMBER, IEEE

Abstract—**This correspondence discusses some possible limitations on the application of a new algorithm for recursive linear estimation. An algorithm proposed by Kailath using Chandrasekhar equations to calculate the Kalman gain matrix has been shown to be computationally advantageous in some special cases. This correspondence considers the case of controllable and observable linear systems and shows that the conditions required for the algorithm to be advantageous form a closed set of Lebesgue measure zero.**

Consider the following controllable and observable autonomous linear system:

$$(L) \qquad \dot{x} = Fx + Gw, \qquad (F, GQ^{1/2}) \text{ controllable}$$
$$y = Hx + v, \qquad (F,H) \text{ observable}$$

where $x \in R^n$, $w \in R^m$, and $v \in R^p$. The initial state $x(0)$ is a normally distributed random variable with mean μ and covariance P_0. The noise processes w and v are Gaussian white noise processes with mean zero and covariances Q and R, respectively. The variables $x(0)$, w, and v are assumed to be independent. If we denote by $\hat{x}(t)$ the conditional mean of $x(t)$ given the observations $y(s)$, $0 \le s \le t$, and by P the covariance matrix of the process $x(t) - \hat{x}(t)$, then we can obtain the following equations of the classical Kalman filter:

$$\dot{\hat{x}} = F\hat{x} + K(t)(y - H\hat{x}), \qquad \hat{x}(0) = \mu \tag{1}$$

$$\dot{P} = FP + PF^T + GQG^T - PH^TR^{-1}HP, \qquad P(0) = P_0 \tag{2}$$

where $K(t) = P(t)H^TR^{-1}$.

We shall use the following properties of these equations [10]. Under the hypotheses of controllability and observability, the covariance matrix $P(t)$ converges to a unique (independent of $P(0)$) positive definite limit $\bar{P}$. Furthermore, the differential equation

$$\dot{\hat{x}} = F\hat{x} + K(t)(y - H\hat{x}) = (F - K(t)H)\hat{x} + K(t)y \tag{3}$$

Manuscript received June 3, 1974; revised October 21, 1974.
The author was with NASA/Goddard Space Flight Center, Network Computing Branch, Greenbelt, Md. He is now with The Analytic Sciences Corporation, Reading, Mass. 01867.

is asymptotically stable. Therefore, if $P(0) = \bar{P}$, then the real parts of the eigenvalues of the matrix

$$F - \bar{K}H = F - \bar{P}H^TR^{-1}H \tag{4}$$

are all negative.

It is evident that the calculation of $K(t)$ requires the calculation of $P(t)$ and that the calculation of $P(t)$ requires the solution of $n(n+1)/2$ simultaneous nonlinear differential equations. If, for example, observations are scalars ($p = 1$), then $K(t)$ contains only n elements, and it would be nice to find a method for calculating $K(t)$ directly if $P(t)$ were not of interest. In [6] Kailath presents an algorithm for calculating $K(t)$, the gain matrix, without calculating P. This algorithm uses differential equations that Kailath describes as being of the Chandrasekhar type, which appear in certain astrophysical problems. In certain circumstances this algorithm is shown to be computationally advantageous. It is shown that if the matrix

$$D = FP(0) + P(0)F^T + GQG^T - P(0)H^TR^{-1}HP(0) \tag{5}$$

has rank α ($\le n$) then the gain matrix $K(t)$ can be computed by solving $n(p + \alpha)$ nonlinear differential equations instead of $n(n+1)/2$. Thus, if $p + \alpha < (n+1)/2$, the algorithm may be useful. The question that naturally arises is the following: which choices of $P(0)$ yield matrices D with low rank? Since the determinant function is a polynomial, it follows that $\det D$ either vanishes for all P_0 or else it vanishes only on a set that is closed in the usual topology and has measure zero with respect to Lebesgue measure.

For controllable and observable systems it is the purpose of this correspondence to show that $\det D$ does not vanish for all P_0. Thus, for almost all P_0, the matrix D has rank n. Thus, in the general case, there is no evident computational advantage in terms of the number of equations to be solved since $n(p + n) > n(n+1)/2$.

Some hypothesis on the coefficient matrices must be made to obtain a result like this. The following well-known facts [5, p. 92] from elementary linear algebra show that for a certain class of system matrices, *all* initial state covariance matrices yield a matrix D with rank $\alpha < n$. Let $\rho(\cdot)$ be the rank function defined on the set of $n \times n$ matrices. The rank function satisfies the inequalities $\rho(AB) \le \min(\rho(A),\rho(B))$ and $\rho(A + B) \le \rho(A) + \rho(B)$.

Using these inequalities, we can obtain an upper bound on the rank of D independent of $P(0)$

$$\rho(D) \le 2\rho(F) + \rho(G) + \rho(H).$$

Consequently, it is easy to generate examples of a system in which $\alpha < n$ for all $P(0)$. The point of the following theorem is that under the hypothesis of controllability and observability, which is the hypothesis of most interest in recursive estimation, this singular behavior of mapping a set of positive measure to a set of measure zero cannot occur even if the system matrices all have rank less than n. What is being proved here is not that the set of matrices D having rank less than n is of measure zero, but that the *inverse image* under the mapping defined by (5) of this set of matrices has measure zero under appropriate hypotheses.

View the formula for D as a mapping from the space of symmetric matrices to itself. This space is isomorphic to $R^{n(n+1)/2}$. Call this mapping S. Thus we have

$$D = S(P), \qquad D: R^{n(n+1)/2} \to R^{n(n+1)/2}.$$