EFFICIENT RECURSIVE ESTIMATION OF THE PARAMETERS OF A RADAR OR RADIO ASTRONOMY TARGET

by

D. J. Sakrison

The research herein was supported by the National Aeronautics and Space Administration under Grant NsG-354(S-2).

December 18, 1964

ABSTRACT

In radar or radio astronomy we observe a signal whose covariance function depends on some target parameters of interest. We consider here the problem of estimating the values of these parameters from our observation of the signal. One possible procedure is to use the method of maximum likelihood estimation. This method has the advantage that, as the duration of the observation interval becomes long, the mean square error in the maximum likelihood estimate approaches the minimum given by the Cramér-Rao bound. However, the maximum likelihood estimate is usually difficult to compute. We present here a recursive estimation procedure which divides the observation interval up into subintervals of short length: on each subinterval the signal is processed quadratically and the resulting calculation used to improve our estimate. This method has many computational advantages and, under certain conditions, we can show that the error in the resulting sequence of estimates approaches the Cramér-Rao bound.

We begin by giving brief consideration to the problem of determining the functional dependence of the covariance function of the received signal on the target parameters. We then present expressions for the terms that appear in the Cramér-Rao inequality. Lastly, we describe the recursive estimation method and state conditions under which it is applicable.

I. INTRODUCTION

In radio or radar astronomy, we observe a signal emitted or scattered from an astronomical body or target with the objective of gaining information concerning the nature of the body. If the observed signal is emitted by the body, the source of the emission is at the atomic level; if the observed signal is a radar signal scattered from the body, the signal is scattered from a very large number of specular points. In either case, the resultant signal is the sum of a large number of incoherent sinusoidal contributions and hence will have zero mean. Further, the central limit theorem is applicable and the signal should be characterized to a good approximation by a gaussian random process. The statistical behavior of such a process is completely described by its covariance function

$$\phi_s(t_1, t_2) = E\{(S_{t_1} - \overline{S}_{t_1})(S_{t_2} - \overline{S}_{t_2})\} \tag{1.1}$$

$$\overline{S}_t = E\{S_t\} = 0$$

and thus all information concerning the body that can be gained by observation of the signal is contained in $\phi_s(t_1, t_2)$.

We consider here the problem of estimating the numerical value of those parameters of the body which may be of interest; we assume the estimate is to be based upon observation of $W(t) = S(t) + N(t)$; the signal corrupted by white gaussian noise; we take the number of parameters of interest to be M, and denote generic values of these parameters by an M-dimensional vector.

$$\underset{\sim}{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m) .$$

The covariance function associated with the observed signal will depend on which values of the parameters actually pertain to the body, and we will denote this functional dependence by writing

$$\phi_s(t_1, t_2) = \phi_s(t_1, t_2, \underset{\sim}{\alpha}) \tag{1.2}$$

Given the problem of trying to estimate some parameter $\underset{\sim}{\alpha}$ in the absence of any a priori statistics, the usual solution is to use the maximum likelihood estimate [1]. As shown by Price [2], the maximum likelihood estimate of $\underset{\sim}{\alpha}$ would be that value of $\underset{\sim}{\alpha}$ which maximizes the quantity

$$\ell[W(t);\underset{\sim}{\alpha}] = \frac{1}{2N} I_{T_0}(\underset{\sim}{\alpha}) - \frac{1}{2} \sum_{n=1}^{\infty} \ell n[1 + \lambda_n(\underset{\sim}{\alpha})/N] \tag{1.3}$$

in which

$$I_{T_0}(\underset{\sim}{\alpha}) = \int_0^{T_0}\int_0^{T_0} h(t, s, \underset{\sim}{\alpha})\, W(t)\, W(s)\, dtds \tag{1.4}$$

and $h(t, s, \underset{\sim}{\alpha})$ is the solution of the integral equation

$$Nh(t, s, \alpha) + \int_0^{T_0} \phi_s(t, r, \underset{\sim}{\alpha})\, h(r, s, \underset{\sim}{\alpha})\, dr = \phi_s(t, s, \underset{\sim}{\alpha}) \tag{1.5}$$

the $\lambda_n(\alpha)$ are the eigenvalues associated with $\phi_s(t, s, \alpha)$, $t, s \in [0, T]$, N denotes the magnitude of the noise spectral density, and T_0 the <u>total</u> observation time. Price also shows [2] that, under the usual weak signal to noise ratio conditions prevailing in radio or radar astronomy, the quantity $I_{T_0}(\underset{\sim}{\alpha})$ is given approximately by

$$I_{T_0}(\underset{\sim}{\alpha}) \approx \int_0^{T_0}\int_0^{T_0} \phi_s(t, s, \underset{\sim}{\alpha})\, W(t)\, W(s)\, dtds \tag{1.6}$$

In the radio astronomy situation, the process observed might be a stationary process and

$$\phi_s(t, s, \underset{\sim}{\alpha}) = \phi_s(t - s, \underset{\sim}{\alpha}) .$$

Using the symmetry of $I_{T_0}(\underset{\sim}{\alpha})$, a change in the variables of integration will yield

$$I_{T_0}(\underset{\sim}{\alpha}) = 2\int_0^T d\tau\,\phi_s(\tau,\underset{\sim}{\alpha})\int_0^{T-\tau} ds\,W(s+\tau)\,W(s) \qquad (1.7)$$

In the radar astronomy case, if the transmission is a sequence of Q sinusoidal transmissions, each of duration T_s, and T_s is large compared to the delay spread of the target, then $I_{T_0}(\underset{\sim}{\alpha})$ can be expressed approximately as a sum of Q integrals of the form of Eq. (1.7).

The use of the maximum likelihood estimation procedure has the advantage that it is asymptotically efficient; that is, as T_0 becomes large the covariance matrix of the errors will approach the minimum given by the Cramér-Rao bound [1] . However, the computational problems associated with this method can be almost prohibitive. The usual situation is one in which a long total observation time (large T_0) is required to obtain a reliable estimate. In this case the data handling and processing problem is enormous (typical signal bandwidths may be 100 c. p. s. or greater and observation times upwards of several hours). Even if $I_{T_0}(\underset{\sim}{\alpha})$ can be expressed in the form of Eq. (1.7) and one-bit correlation methods [3, 4] used to calculate the time correlation function of W, the computational problems associated with evaluating $I_{T_0}(\underset{\sim}{\alpha})$ and implementing a search or gradient seeking procedure for maximizing $\ell(W;\underset{\sim}{\alpha})$ are not trivial.

To solve the computational problem, we propose a recursive estimation method. That is, we will divide the total observation interval of duration T_0 into a large number of short subintervals, each of duration T. The duration T can be picked for compatability with computational facilities or on some other basis of convenience, as long as it is long with respect to the reciprocal of the bandwidth of S(t). We start with some arbitrary estimate $\underset{\sim}{\alpha}_1$. During the reception of W(t) in the first subinterval of length T, the signal W(t) is processed <u>once</u> by a quadratic processor and a correction term based on this calculation is added to $\underset{\sim}{\alpha}_1$ to yield $\underset{\sim}{\alpha}_2$.

This process continues: at the start of the n-th subinterval of duration T we have the estimate $\underset{\sim}{\alpha}_n$ which we update by quadratically processing $W(t)$ during the n-th reception subinterval and adding a correction term to $\underset{\sim}{\alpha}_n$. The computational advantage of such a method lies in the fact that no more than T seconds of data is ever handled by the computer and the only quantity that needs to be stored from one interval to the next is $\underset{\sim}{\alpha}_n$, the current estimate.

In Sec. 4, we will present a specific such recursive method, and state conditions which are sufficient to guarantee that $\underset{\sim}{\alpha}_n$ converges to the true value of the target parameters. This method will be asymptotically efficient; that is, when $\underset{\sim}{\alpha}_n$ does converge, the resulting error covariance matrix tends to the bound given by the Cramér-Rao inequality [1] as T_0 becomes large. Before discussing this method, we briefly discuss in Sec. 2 the relation between $\phi_s(t, s, \underset{\sim}{\alpha})$ and the properties of the target in the radar-astronomy situation. Section 3 presents explicit closed form expressions for the matrix appearing in the Cramér-Rao inequality and several quantities relevant to the discussion of Sec. 4. Section 5 concludes with an example.

II. RELATION OF THE SIGNAL CORRELATION FUNCTION TO THE PARAMETERS OF A RADAR TARGET

In the radio astronomy case there is nothing that we can say in general relating the parameters of interest and the covariance function of the radiated signal; in each case the relation between the parameters and the physical origin of the radiation must be considered.

In the case of a rigid radar target, a useful concept relating the scattered radiation and the target parameters is that of the scattering function discussed by Green [5]. Consider the surface of the target divided up on the basis of range and relative velocity with respect to the radar antenna. The scattering function is denoted by $\sigma(\tau, \omega)$ and $\sigma(\tau, \omega)\, d\tau\, d\omega$ denotes the effective area of radar cross section of those portions of the target located at delays between $\tau - d\tau$ and τ and doppler shifts between $\omega - d\omega$ and ω. The average power returned from those portions of the target lying in these delay and doppler zones is thus proportional to $\sigma(\tau, \omega)\, d\tau\, d\omega$.

The dependence of the target scattering function upon the target parameters is usually direct, and enables one to find the functional relationship $\sigma(\tau, \sigma, \underset{\sim}{\alpha})$. For example, Green [5] has derived an expression for the target scattering function of a rough rotating sphere which directly expresses its dependence upon the rotational velocity and angular scattering function of the sphere.

Inasmuch as our results are based on the dependence of the covariance function on the target parameters, it would be useful to have a relation between $\phi_s(t_1, t_2)$ and $\sigma(\tau, \omega)$,

Consider transmitting the narrow band signal

$$x(t) = \mathrm{Re}\{\chi(t) \exp(j\Omega t)\} \tag{2.1}$$

in which $\chi(t)$ is a complex valued lo-pass signal. The signal, $S_{\tau\omega}(t)$, returned from that position of the target at delay τ and doppler shift ω, is then the real part of

$$\zeta_{\tau,\omega}(t) = K\chi(t - \tau)\, r(t - \tau, \tau, \omega)\, \sigma^{1/2}(\tau, \omega) \exp[\, j(\Omega+\omega)(t-\tau)\,] \tag{2.2}$$

in which the constant K depends on the antenna gain and range of the target. The random time-varying complex-valued reflectivity coefficient r describes the variations in the area and number of specular points at τ, ω caused either by rotation or by local motion if the scattering is taking place off a gaseous atmosphere.

We assume r is normalized such that

$$E\{\,|r|^2\} = 1$$

Although the variation in delay over a region between $\tau - d\tau$ and τ is negligible with respect to the modulation, there will still be fluctuations large with respect to a carrier cycle. This would imply first that

$$E\{r(t, \omega, \tau)\} \equiv 0\,. \tag{2.3}$$

Secondly, one would expect that if a steady sinusoidal signal were scattered from the region, the covariance function of the scattered signal should be independent of the phase of the transmitted signal. It can be shown, in a straightforward but laborious manner, that this implies that

$$E\{r(t_1, \tau, \omega)\, r(t_2, \tau, \omega)\} \equiv 0 \tag{2.4}$$

Lastly, these same considerations imply that $r(t, \tau, \omega)$ will be uncorrelated at regions of different delay and doppler shifts; thus, applying Eq. (2.3) and (2.4), we have

$$E\{r(t_1, \tau_1, \omega_1)\, r^*(t_2, \tau_2, \omega_2)\} = \delta(\tau_1 - \tau_2)\,\delta(\omega_1 - \omega_2)\,\rho(t_1 - t_2, \tau_1, \omega_1) \tag{2.5}$$

$$E\{r(t_1, \tau_1, \omega_1)\, r(t_2, \tau_2, \omega_2)\} = 0 \tag{2.6}$$

Returning to Eq. (2.2), we have that the total returned signal is given by $S(t) = \text{Re}\{\zeta(t)\}$, in which

$$\zeta(t) = K \iint d\tau\, d\omega\, \chi(t-\tau)\, r(t-\tau, \tau, \omega)\, \sigma^{1/2}(\tau, \omega) \exp[j(\Omega+\omega)(t-\tau)] \tag{2.7}$$

the integration being over all values of delay and doppler shift associated with the target. Using Eq. (2.7) to write out the quantities

$$S(t_1)\, S(t_2) = [\text{Re}\, \zeta(t_1)]\, [\text{Re}\, \zeta(t_2)]$$

and

$$\text{Re}\, [\zeta(t_1)\, \zeta^*(t_2)]$$

taking expected values, and using Eqs. (2.5) and (2.6) yield

$$\phi_s(t_1, t_2) = E\{S_{t_1}, S_{t_2}\} = (1/2)\text{Re}\, E\{\zeta_{t_1} \zeta^*_{t_2}\} =$$

$$= (1/2) K^2 \mathrm{Re} \iint d\tau\, d\omega\, \chi(t_1 - \tau) \chi^*(t_2 - \tau) \rho(t_1 - t_2, \tau, \omega)$$

$$\times \sigma(\tau, \omega) \exp[j(\Omega + \omega)(t_1 - t_2)] \qquad (2.8)$$

Lastly, if we assume that $\rho(t_1-t_2, \tau, \omega)$ is independent of τ and ω, we have

$$\phi_s(t_1, t_2) = \frac{K^2}{2} \mathrm{Re}\, \rho(t_1 - t_2) \int d\tau\, \chi(t_1 - \tau) \chi^*(t_2 - \tau)$$

$$\times \int d\omega\, \sigma(\tau, \omega) \exp[j(\Omega + \omega)(t_1 - t_2)] \qquad (2.9)$$

III. EXPRESSIONS RELATED TO THE LIKELIHOOD FUNCTION AND THE CRAMÉR-RAO BOUND

In this section we present an expression for the matrix appearing in the Cramér-Rao inequality [1, 6]. This inequality is lower bound on the error that can be achieved in estimating a linear combination of the components of $\underset{\sim}{\alpha}$ based on some observation related to $\underset{\sim}{\alpha}$. Let $\hat{\underset{\sim}{\alpha}}$ be any unbiased estimator of $\underset{\sim}{\alpha}$, $\underset{\sim}{\theta}$ be the "true" value of $\underset{\sim}{\alpha}$, and $\underset{\sim}{c}$ an arbitrary vector. Then the Cramér-Rao inequality is

$$E\left\{\left[\sum_{j=1}^{M} c_j(\hat{\alpha}_j - \theta_j)\right]^2\right\} = \underset{\sim}{c}' R_{\hat{\alpha}} \underset{\sim}{c} \geq \underset{\sim}{c}' B^{-1} \underset{\sim}{c} \qquad (3.1)$$

in which prime denotes transpose, $R_{\hat{\alpha}}$ is the matrix whose ij-th element is

$$E\{(\alpha_i - \theta_i)(\alpha_j - \theta_j)\}$$

and B is the matrix whose ij-th element is

$$b_{ij} = E\left\{\frac{\partial \ell}{\partial \alpha_i} \frac{\partial \ell}{\partial \alpha_j}\bigg|_{\underset{\sim}{\alpha} = \underset{\sim}{\theta}}\right\} \qquad (3.2)$$

ℓ denoting the log-likelihood function* of $\underset{\sim}{\alpha}$ based on the observation made. The quantities $\partial \ell / \partial \alpha_i$ and b_{ij} are thus of direct interest, and we now present closed form expressions for these pertinent quantities.

Consider observing the random process $W(t) = S(t) + N(t)$, the signal of interest plus white gaussian noise, on the time interval $[0, T]$. Let $\psi_n(t, \underset{\sim}{\alpha})$ and $\lambda_n(\underset{\sim}{\alpha})$ be the normalized eigenfunctions and eigenvalues of the integral equation

$$\int_0^T ds\ \phi_s(t, s, \underset{\sim}{\alpha})\, \psi_n(s, \underset{\sim}{\alpha}) = \lambda_n(\underset{\sim}{\alpha})\, \psi_n(t, \alpha) \qquad 0 \leq t \leq T \tag{3.3}$$

If $\underset{\sim}{\alpha}$ were the value of the parameter vector actually associated with the process $S(t)$, then the process $W(t)$ would be described by [8]

$$W(t) = \underset{n \to \infty}{\text{l.i.m.}} \sum_{k=1}^{n} W_k \psi_k(t, \underset{\sim}{\alpha}) \tag{3.4}$$

in which

$$W_k = \int_0^T W(t)\, \psi_k(t, \underset{\sim}{\alpha}) dt \tag{3.5}$$

The coefficients W_k are all zero mean independent gaussian random variables with variance $\lambda_n(\underset{\sim}{\alpha}) + N_0$, N_0 being the (two sided) spectral density of the additive white noise. The likelihood ratio of $\underset{\sim}{\alpha}$ based on the observation $W(t)$, $t \in [0, T]$, can then be taken as the limit

---

* To be strictly correct, the quantity ℓ appearing in this expression should be the Radon-Nikodyn derivative. If the correlation function of W is continuous our definition of ℓ as defined in Eq. (3.6) would be equal with probability one to the Radon-Nikodyn derivative [7]. In actual fact, the noise N is not white but merely very broad band so this condition would be met. However, for simplicity of the resultant expressions, we will use an expression for ℓ [Eq. (3.9)] that results from letting the bandwidth of N become infinite.

$$\Lambda(W;\underset{\sim}{\alpha}) = \lim_{n \to \infty} \frac{p(W_1, \ldots, W_n;\underset{\sim}{\alpha})}{p(W_1, \ldots, W_n;0)} \tag{3.6}$$

in which $p(W_1, \ldots, W_n;\underset{\sim}{\alpha})$ is the density function associated with $W_1, \ldots, W_n$ under the hypothesis that the process $S(t)$ is characterized by the parameter value $\underset{\sim}{\alpha}$ and $p(W_1, \ldots, W_n;0)$ is the same density function under the hypothesis that the process $W(t)$ is the noise alone. It is shown [9] that the natural log of this quantity is given by

$$\ell(W;\underset{\sim}{\alpha}) = (1/2N) \sum_{n=1}^{\infty} W_n^2[\lambda_n(\underset{\sim}{\alpha})/(\lambda_n(\underset{\sim}{\alpha}) + N)] - (1/2)\ln[\prod_{n=1}^{\infty}(\lambda_n(\underset{\sim}{\alpha}) + N)] \tag{3.7}$$

If $h(t, s, \underset{\sim}{\alpha})$ is the solution of the integral equation

$$Nh(t, s, \underset{\sim}{\alpha}) + \int_0^T dr\, \phi_s(t, r, \underset{\sim}{\alpha})\, h(r, s, \underset{\sim}{\alpha})\, dr = \phi_s(t, s, \underset{\sim}{\alpha}) \qquad t, s, \epsilon\ [0, T] \tag{3.8}$$

then the first term in Eq. (3.7) may be expressed in closed form, and

$$\ell(W;\underset{\sim}{\alpha}) = \frac{1}{2N}\int_0^T\int_0^T h(t, s, \underset{\sim}{\alpha})\, W(t)\, W(s)\, dtds$$

$$- (1/2) \ln [\prod_{n=1}^{\infty} (\lambda_n(\underset{\sim}{\alpha}) + N)] \tag{3.9}$$

To develop further expressions, we will need to interchange certain operations such as integration and partial differentiation. These interchanges are justified only if $\phi_s(t, s, \underset{\sim}{\alpha})$ is suitably well behaved; thus we now state conditions that will be sufficient to justify these operations.

Condition 1: $W(t) = S(t) + N(t)$ in which S and N are independent zero mean gaussian processes; the noise $N(t)$ is white with spectral density N.

Condition 2: Let a superscript denote partial differentiation with respect to the corresponding component of $\underset{\sim}{\alpha}$. The functions $\phi_s(t, s, \underset{\sim}{\alpha})$, $h(t, s, \underset{\sim}{\alpha})$, and $h^i(t, s, \underset{\sim}{\alpha})$ are assumed to be continuous in t and s on $[0, T] \times [0, T]$ and the functions $\phi^i(t, s, \underset{\sim}{\alpha})$ and $h^{ij}(t, s, \underset{\sim}{\alpha})$ assumed to be integrable square over $[0, T] \times [0, T]$ for $i, j = 1, 2, \ldots, M$.

Let us now take the partial $\partial / \partial \alpha_i$ of both sides of Eq. (3.9) and interchange differentiation and integration on the right hand side (the above conditions justify this interchange [10]) to obtain

$$\partial \ell / \partial \alpha_i = \frac{1}{2N} \int_0^T \int_0^T h^i(t, s, \underset{\sim}{\alpha})\, W(t)\, W(s)\, dtds$$

$$- (1/2) \sum_{n=1}^{\infty} \frac{\partial \lambda_n(\alpha)}{\partial \alpha_i} / (\lambda_n(\underset{\sim}{\alpha}) + N) \tag{3.10}$$

Starting with the integral equation (3.3), it is possible to find an equivalent closed form expression for the sum in Eq. (3.10) [10]; the result is

$$\partial \ell / \partial \alpha_i = \frac{1}{2N} \int_0^T \int_0^T h^i(t, s, \underset{\sim}{\alpha}) \left[W(t) W(s) - \phi_w(t, s, \underset{\sim}{\alpha}) \right] dtds \tag{3.11}$$

in which $\phi_w(t, s, \underset{\sim}{\alpha}) = \phi_s(t, s, \underset{\sim}{\alpha}) + N\delta(t-s)$. From this expression and Eq. (3.2) one can obtain the b_{ij} appearing in the Cramér-Rao bound. The resulting expression reduces after some manipulation [10] to

$$b_{ij} = \frac{1}{2N} \int_0^T \int_0^T h^i(t, s, \underset{\sim}{\theta})\, \phi_s^j\, (t, s, \underset{\sim}{\theta})\, dtds \tag{3.12}$$

These quantities, when used in inequality (3.1), bound the performance of an unbiased estimate of $\underset{\sim}{\alpha}$ based on a single observation of $W(t)$ of duration T. It can be shown directly that the corresponding bound for an estimate based on n statistically independent such observations is simply the right hand side of Eq. (3.1) divided by n.

Reflection upon Eq. (3.11) can suggest an iterative method for estimating $\underset{\sim}{\theta}$. Take the expected value of both sides of this equation, noting that condition (2) allows the expectation and integration to be interchanged.

$$E\left\{\frac{\partial \ell}{\partial \alpha_i}\right\} = \frac{1}{2N}\int_0^T\int_0^T h^i(t, s, \underset{\sim}{\alpha}) \; [\phi_W(t, s, \underset{\sim}{\theta}) - \phi_W(t, s, \underset{\sim}{\alpha})]\,dtds \tag{3.13}$$

For convenience we denote $E\{\partial \ell/\partial \alpha_i\}$ by $m_i(\underset{\sim}{\alpha})$ and the M-dimensional vector whose i-th component is $m_i(\underset{\sim}{\alpha})$ by $\underset{\sim}{m}(\underset{\sim}{\alpha})$. Note that $\underset{\sim}{m}(\underset{\sim}{\alpha})$ is equal to $\underset{\sim}{0}$ for $\underset{\sim}{\alpha} = \underset{\sim}{\theta}$. We will restrict ourselves to situations in which this is the only value of $\underset{\sim}{\alpha}$ satisfying this equation and use a recursive search procedure based on successive observations of duration T of the process W(t) to find the value of $\underset{\sim}{\alpha}$ that sets $\underset{\sim}{m}(\underset{\sim}{\alpha}) = \underset{\sim}{0}$.

IV. A RECURSIVE ESTIMATION METHOD

We now wish to investigate the possibility of recursively estimating $\underset{\sim}{\theta}$, the value of $\underset{\sim}{\alpha}$, that satisfies $m(\underset{\sim}{\alpha}) = \underset{\sim}{0}$, by making successive observations of W(t). Further we wish to do this in such a manner that the resulting sequence of estimates is asymptotically efficient; that is, as n, the number of observations, becomes large, the covariance matrix of the errors, $R_{\hat{\alpha}}$, should approach (1/n)B, the entries of B being given by Eq. (3.12).

To facilitate the discussion let us denote by $\underset{\sim}{Y}_n(\underset{\sim}{\alpha})$ the M-dimensional vector whose i-th component is $\partial \ell/\partial \alpha_i$ evaluated from the observation of W(t) on the n-th time interval of duration T. Note that $\underset{\sim}{Y}_n(\underset{\sim}{\alpha})$ is calculated by the M quadratic operations of of Eq. (3.11), in which the integration is carried out over the n-th observation interval instead of [0, T]. We shall assume that observations of W(t) made on these (disjoint) intervals of time are statistically independent. Now consider finding $\underset{\sim}{\theta}$, the value of $\underset{\sim}{\alpha}$ for which $\underset{\sim}{m}(\underset{\sim}{\alpha})$ is zero; under suitable conditions on $\underset{\sim}{m}(\underset{\sim}{\alpha})$, we might regard $\underset{\sim}{m}(\underset{\sim}{\alpha})$ as the gradient of a unimodal surface. If we were able to make successive observations of $\underset{\sim}{m}(\underset{\sim}{\alpha})$ at different values of $\underset{\sim}{\alpha}$, then we could employ a conventional gradient seeking or hill climbing method to search out $\underset{\sim}{\theta}$. Although we are not able to observe $\underset{\sim}{m}(\underset{\sim}{\alpha})$, we can observe $\underset{\sim}{Y}_n(\underset{\sim}{\alpha})$, $n = 1, 2, \ldots,$ where $E\{\underset{\sim}{Y}_n(\underset{\sim}{\alpha})\} = \underset{\sim}{m}(\underset{\sim}{\alpha})$. If we carry out the usual gradient seeking procedure using the $\underset{\sim}{Y}_n$ and weight of the sequence of resulting corrections by a "gain" that decreases as the observation number increases, the fluctuations in $\underset{\sim}{Y}_n$ tend to be cancelled out and, under suitable conditions, the resulting sequence of estimates tends to converge to $\underset{\sim}{\theta}$.

To make this explicit, consider the sequence of estimates $\underset{\sim}{\alpha}_n$, $n = 1, 2, \ldots,$ in which $\underset{\sim}{\alpha}_1$ is chosen arbitrarily and the remainder of the estimates are determined by the recursion equation

$$\underset{\sim}{\alpha}_{n+1} = \underset{\sim}{\alpha}_n + (A/n)\, \underset{\sim}{Y}_n(\underset{\sim}{\alpha}) \tag{4.1}$$

Such methods are known as Stochastic Approximation methods; they have been studied for some time and the convergence of $\underset{\sim}{\alpha}_n$ to $\underset{\sim}{\theta}$ has been proven under a variety of conditions [11, 12, 13, 14]. The difficulty with existing methods is that the covariance matrix of the errors R_{α_n}, depends very critically upon the gain constant A. A poor choice of A can result in a mean square error

$$E\{\| \underset{\sim}{\alpha}_n - \underset{\sim}{\theta} \|^2,\}$$

which is much larger than that given by the Cramér-Rao bound [14]. Unfortunately this dependence upon A also depends upon the unknown value of $\underset{\sim}{\theta}$, hence a good a priori choice of A is not possible.

To correct this situation, we consider a stochastic equivalent of a Newton-Raphson procedure in which the correction terms are weighted by the inverse of the matrix of second partials of the surface.

Specifically, let

$$\underset{\sim}{\alpha}_{n+1} = \underset{\sim}{\alpha}_n + (1/n)G^{-1}(\underset{\sim}{\alpha}_n)\underset{\sim}{Y}_n(\underset{\sim}{\alpha}_n) \quad (3.2)$$

in which the ij-th entry of the matrix $G(\underset{\sim}{\alpha})$ is given by

$$g_{ij}(\underset{\sim}{\alpha}) = \frac{1}{2N} \int_0^T \int_0^T h^i(t, s, \underset{\sim}{\alpha})\, \phi^j(t, s, \underset{\sim}{\alpha})\, dtds \quad (3.3)$$

It can be shown directly from Eq. (3.13) using condition 2 that

$$-\frac{\partial}{\partial \alpha_j} m_i(\underset{\sim}{\alpha}) \Big|_{\underset{\sim}{\alpha}=\theta} = g_{ij}(\underset{\sim}{\theta}) \quad (3.4)$$

thus as $\underset{\sim}{a} \to \underset{\sim}{\theta}$ the matrix $G(\underset{\sim}{\alpha})$ approaches the matrix of second partials that appear in the Newton-Raphson method. Note that $g_{ij}(\underset{\sim}{\theta})$ is equal to the g_{ij} of the Cramér-Rao bound.

Under suitable conditions, it can be shown that the sequence of estimates approaches $\underset{\sim}{\theta}$, and the covariance matrix of the errors approaches (1/n) G for large n. We now state conditions which are sufficient to guarantee this.

Condition 3: Observations of $W(t)$ made on the disjoint time intervals of duration T are statistically independent and identically distributed.

Condition 4: $\underset{\sim}{\theta}$ is known a priori to lie in the interior of some bounded set A (in the M-dimensional space of parameter values) and the sequence of estimates generated by Eq. (3.2) is constrained to lie in this set.

Condition 5: $G(\underset{\sim}{\alpha})$ is invertible for all $\underset{\sim}{\alpha} \in A$.

Condition 6: The quantities

$$\frac{\partial^2}{\partial \alpha_k \partial \alpha_\ell} \sum_{j=1}^{M} g_{ij}(\underset{\sim}{\alpha})\, m_j(\underset{\sim}{\alpha})$$

and

$$\frac{\partial}{\partial \alpha_\ell} \sum_{i,j,k=1}^{M} b_{ik}(\underset{\sim}{\alpha})\, g_{ji}(\underset{\sim}{\alpha})\, g_{jk}(\underset{\sim}{\alpha})$$

in which $b_{ij}(\alpha) = E\left\{\frac{\partial \ell}{\partial \alpha_i} \frac{\partial \ell}{\partial \alpha_j}\right\}$ are bounded for all $\underset{\sim}{\alpha} \in A$.

Condition 7: There exist a K_0 and K_0', $0 < K_0 \leq K_0' < \infty$ such that

$$K_0 \;\|\underset{\sim}{\alpha} - \underset{\sim}{\theta}\|^2 \leq -(\underset{\sim}{\alpha} - \underset{\sim}{\theta})' \underset{\sim}{G}(\underset{\sim}{\alpha}) \underset{\sim}{m}(\underset{\sim}{\alpha}) \leq K_0' \;\|\underset{\sim}{\alpha} - \underset{\sim}{\theta}\|^2$$

for all $\underset{\sim}{\alpha} \in A$.

Condition 3 can be assumed to hold in practice as long as the bandwidth of $S(t)$ is large compared to $1/T$. The remainder of Conditions 2-6 are of the nature of regularity conditions and will usually be satisfied in practice. Condition 7 is restrictive and will seriously limit the situations to which our method is applicable. Condition 7 is required to guarantee that the equation $\underset{\sim}{m}(\underset{\sim}{\alpha}) = \underset{\sim}{0}$ have as its only solution $\underset{\sim}{\alpha} = \underset{\sim}{\theta}$; i.e., that the "surface" whose maximum we are locating be unimodal.

We have shown the following result:

Theorem: Conditions 1-7 imply that the sequence of estimates $\underset{\sim}{\alpha}_n$ generated by Eq. (3.2) is asymptotically efficient, i.e.

$$E\left\{(\alpha_{n,i} - \theta_i)(\alpha_{n,j} - \theta_j)\right\} = (1/n) g_{ij} + K n^{-(1+\gamma)} \qquad K < \infty,\ \gamma > 0$$

This statement is proved in [10]; the proof is long and involved and we do not give it here.

In the closing section we present an example to indicate the method's applicability to a problem of practical interest and the scope of the limitation imposed by Condition 7.

V. AN EXAMPLE

We now consider the case in which $S(t)$ is a gaussian process whose correlation function is of the form

$$\phi_s(t, s, \underset{\sim}{\alpha}) = A \exp[-\gamma|t-s|] \cos \omega(t-s) \tag{5.1}$$

in which A, γ, and ω are unknown and to be estimated. Reflection upon Eqs. (3.2) and (3.4) reveals that in order for Eq. (3.2) to be correct dimensionally the α_i's must all have the same dimension. We will take them to be dimensionless and rewrite Eq. (5.1) as

$$\phi_s(t, s, \underset{\sim}{\alpha}) = \alpha_1 A_r \exp[-\gamma_r \alpha_2 |t-s|] \cos \alpha_3 \gamma_r (t-s) \tag{5.2}$$

in which A_r and γ_r are arbitrary reference values. In order that condition 3 be sati fied we will need to pick T such that

$$\alpha_2 \gamma << T$$

for all values of α_2 that are regarded a priori as possible.

For $\phi_s(t, s, \alpha)$ given by Eq. (5.2) the function $h(t, s, \alpha)$ can be found directly [15,16]. However, the computations involved are somewhat laborious, and we will make an approximation. In the usual case in radio astronomy, the signal S would be grossly weaker than the noise; i.e.,

$$\frac{\alpha_1 A_r}{\alpha_2 \gamma_r} << N \tag{5.3}$$

In this situation, expansion of $h(t, s, \underset{\sim}{\alpha})$ in a Von Neumann series indicates that under the conditions of inequality (5.3)

$$h(t, s, \underset{\sim}{\alpha}) \simeq (1/N) \ \phi_s(t, s, \underset{\sim}{\alpha}) \ . \tag{5.4}$$

For convenient normalization we shall set

$$A_r/\gamma_r = N$$

(it can be shown 10 that condition 7 is independent of the scale factor chosen) so that we are interested in those situations in which

$$\alpha_1/\alpha_2 << 1 \tag{5.5}$$

Further, most cases of interest will be those in which S(t) is narrow band; hence we assume

$$\alpha_2/\alpha_3 << 1 \tag{5.6}$$

Using the approximation of Eq. (5.4) and using inequalities (5.5) and (5.6) to simplify the resulting expressions, we have calculated for $G(\alpha)$

$$G(\alpha) \simeq \frac{8N\alpha_2}{\gamma_r T(\alpha_1)^2} \begin{bmatrix} \alpha_1^2 & \alpha_1\alpha_2 & 0 \\ \alpha_1\alpha_2 & 2(\alpha_2)^2 & 0 \\ 0 & 0 & \alpha_2^2 \end{bmatrix} \tag{5.7}$$

and for $\underset{\sim}{m}(\underset{\sim}{\alpha})$

$$m_1(\underset{\sim}{\alpha}) = F\{2\alpha_2[(\alpha_2 + \theta_2)^2 + (\alpha_3 - \theta_3)^2][(\alpha_2 + \theta_2)^2(\alpha_1 - \theta) - \theta_1(\alpha_2 + \theta_2)(\alpha_2 - \theta_2) + \alpha_2(\alpha_3 - \theta_3)^2]\} \tag{5.8}$$

$$m_2(\underset{\sim}{\alpha}) = F\{2\alpha_1\theta_1(\alpha_2 + \theta_2)^2(\alpha_2 - \theta_2) - \alpha_1(\alpha_1 - \theta_1)(\alpha_2 + \theta_2)^4 + \alpha_1\theta_1(\alpha_2 + \theta_2)^2(\alpha_2 - \theta_2)^2 - \alpha_1(\alpha_3 - \theta_3)^2[\theta_1(2\alpha_2)^2 + 2\alpha_1(\alpha_2 + \theta_2)^2 + \alpha_1(\alpha_3 - \theta_3)^2] \tag{5.9}$$

$$m_3(\underset{\sim}{\alpha}) = F\{2(2\alpha_2)^2 \alpha_1\theta_1(\alpha_2 + \theta_2)(\alpha_3 - \theta_3)\} \tag{5.10}$$

in which

$$F = - \frac{\gamma_r T}{8N(\alpha_2)^2[(\alpha_2 + \theta_2)^2 + (\alpha_3 - \theta_3)^2]} \tag{5.11}$$

From Eqs. (5. 7) -(5. 11) we can calculate the quantity $(\underset{\sim}{\alpha} - \underset{\sim}{\theta})'G(\underset{\sim}{\alpha})\underset{\sim}{m}(\underset{\sim}{\alpha})$. Making use of Eqs. (5. 5) and (5. 6) to obtain a simplified approximation to this quantity, we have

$$\begin{aligned} -(\underset{\sim}{\alpha} - \underset{\sim}{\theta})G(\underset{\sim}{\alpha})\underset{\sim}{m}(\underset{\sim}{\alpha}) \simeq 2\alpha_1\theta_1(\alpha_2)^2(\theta_2)^5 \Big\{ &(\Delta\alpha_2)^2 [4 + 8\Delta\alpha_2 + 5(\Delta\alpha_2)^2 \\ &+ (\Delta\alpha_2)^3] + (\Delta\alpha_3)^2 [8 + (16) + 4\frac{\theta_1}{\alpha_1}\Delta\alpha_1) \ \Delta\alpha_2 \\ &+ (6 + 4\Delta\alpha_1)(\Delta\alpha_2)^2 + (\Delta\alpha_1 - 1)(\Delta\alpha_2)^3] \Big\} \end{aligned} \tag{5.12}$$

in which

$$\Delta\alpha_1 = \frac{\alpha_1 - \theta_1}{\theta_1}, \quad \Delta\alpha_2 = \frac{\alpha_2 - \theta_2}{\theta_2}, \quad \Delta\alpha_3 = \frac{\alpha_3 - \theta_3}{\theta_3}$$

Condition 7 will be satisfied for those values of $\underset{\sim}{\alpha}$ for which the right hand side of Eq. (3.12) is positive; thus, as long as $\underset{\sim}{\alpha}$ is constrained to lie in a set for which this expression is positive, our recursive estimation method may be applied and will be asymptotically efficient. The table below gives a brief list of inequality constraints whose satisfaction implies that $-(\underset{\sim}{\alpha} - \underset{\sim}{\theta})' G(\underset{\sim}{\alpha}) \underset{\sim}{m}(\underset{\sim}{\alpha})$ is positive. The expressions are given in terms of the original A, γ and ω parameters of Eq. (5.1), A_0, γ_0, and ω_0 denoting the true values of A, γ, and ω and $\Delta\omega$ denoting $\omega - \omega_0$.

	No constraint on $\Delta\omega$	
$0 < A \leq 2A_0$	$0 < A \leq 3A_0$	$0 < A \leq 6A_0$
$0.45\gamma_0 \leq \gamma \leq 3.2\gamma_0$	$0.56\gamma_0 \leq \gamma \leq 3.2\gamma_0$	$0.73\gamma_0 \leq \gamma \leq 3.2\gamma_0$
	$0 \leq \lvert\Delta\omega\rvert \leq 4\gamma_0$	
$0 < A \leq 2A_0$	$0 < A \leq 3A_0$	$0 < A \leq 6A_0$
$0.435\gamma_0 \leq \gamma < \infty$	$0.56\gamma_0 \leq \gamma < \infty$	$0.73\gamma_0 \leq \gamma < \infty$
	$0 \leq \lvert\Delta\omega\rvert \leq 2\gamma_0$	
$0 < A \leq 2A_0$	$0 < A \leq 3A_0$	$0 < A \leq 6A_0$
$0.43\gamma_0 \leq \gamma < \infty$	$0.56\gamma_0 \leq \gamma < \infty$	$0.73\gamma_0 \leq \gamma < \infty$
	$0 \leq \lvert\Delta\omega\rvert \leq \gamma_0$	
$0 < A \leq 2A_0$	$0 < A \leq 3A_0$	$0 < A \leq 6A_0$
$0.41\gamma_0 \leq \gamma < \infty$	$0.55\gamma_0 \leq \gamma < \infty$	$0.72\gamma_0 \leq \gamma < \infty$
	$0 \leq \lvert\Delta\omega\rvert \leq (1/2)\gamma_0$	
$0 < A \leq 2A_0$	$0 < A \leq 3A_0$	$0 < A \leq 6A_0$
$0.37\gamma_0 \leq \gamma < \infty$	$0.51\gamma_0 \leq \gamma < \infty$	$0.71\gamma_0 \leq \gamma < \infty$

From the above table, we can see that the method cannot be applied with unqualified success to this example; that is, it is not possible to apply this method when there is no a priori information regarding all three parameters, A, γ, and ω. However, from the above table we also note that if any one of the three parameters is known to lie within a fairly narrow interval, a wide latitude of values is allowed for the other two. Fortunately, in many problems of interest there is a priori information concerning one of the three parameters which allows us to restrict our search in this manner. Thus, although Condition 7 places definite limits on the range of application of our recursive estimation method, there will be problems of interest to which it does apply. For these, its computational simplicity and asymptotic efficiency make it attractive.

REFERENCES

1. Wilks, S.S. (1962). Mathematical Statistics, John Wiley, New York, Chap. 12.

2. Price, R. (May, 1962). Maximum Likelihood Estimation of the Correlation Function of a Threshold Signal, and its Application to the Measurement of the Target Scattering Function in Radar Astronomy. Report 34 G-4, Lincoln Laboratory, Lexington, Mass.

3. Weinrab, S. (Aug. 1963). A Digital Spectral Analysis Technique and its Application to Radio Astronomy. Technical Report 412, Research Laboratory of Electronics, M. I. T., Cambridge 39, Mass.

4. Goldstein, R. (1964). The Analysis of Uncooperative Radar Targets, Technical Report No. 32-658, Jet Propulsion Laboratories, Pasadena, Calif.

5. Green, P. (Dec. 1962). Radar Astronomy Measurement Techniques, Technical Report No. 282, Lincoln Laboratory, Lexington, Mass.

6. Grenander (1950), "Stochastic processes and statistical inference," Arkiv För Matematic, Band 1 nr 17, 247-248.

7. Kadota, (Sept. 1964), "Optimum reception of binary gaussian signals," Bell System Technical Journal.

8. Davenport, W. and Root, W. L. (1958), Random Signals and Noise, McGraw Hill, New York, Chap. 6.

9. Helstrom, C. R. (1960), Statistical Theory of Signal Detection, Pergamon, Oxford, Chap. 11.

10. Sakrison, D. (to appear), Efficient Recursive Estimation; Application to Estimating the Parameters of a Covariance Function, International Journal of Engineering Science.

11. Robbins, H. and Monro, S. (1951), "A stochastic approximation method." Ann. Math. Statist., 22, 400-407.

12. Kiefer, J. and Wolfowitz, J. (1952), "Stochastic estimation of the maximum of a regression function." Ann. Math. Statist., 23, 462-466.

13. Blum, J., (1954), "Multidimensional stochastic approximation methods," Ann. Math. Statist., 23, 462-466.

14. Sacks, J. (1958), "Asymptotic distribution of stochastic approximation procedures," Ann. Math. Statist., 29, 373-405.

15. Zadeh, L., and Ragazzini, J. (1950), "An extension of Wiener's theory of prediction," J. App. Phys., 21, 645-655.

16. Davenport, W. and Root, W., op. cit, 240-242.