N65-35259 [redacted]

UNPUBLISHED PRELIMINARY DATA [redacted]

EFFICIENT RECURSIVE ESTIMATION OF THE PARAMETERS OF A COVARIANCE FUNCTION[1]

34p.

NASA CR 58090 [redacted]

David J. Sakrison
University of California
Berkeley, California

Cat 07

1. Introduction. We consider the situation in which an observer sequentially makes observations, each of T seconds duration, of a zero-mean gaussian (normal) process, W(t). We assume here that successive observations are identically distributed and statistically independent of one another. We consider the case in which the covariance function of the process is of a known form which depends upon the values of M unknown scalar parameters, $\alpha_1, \alpha_2, \cdots, \alpha_M$. The observers objective is to use the known functional form of the covariance function and the sequence of observations to estimate the values of the α-parameters which pertain to the process being observed

In many applications, the amount of data to be processed is extraordinarily large, and the computational aspects of an estimation method become crucially important. A situation which falls within this setting (and indeed was the motivation for this work) is one in radar (or radio) astronomy. Here a signal is periodically scattered (or continuously radiated) from a target such as a planet. Due to the scattering (or radiating) properties of the target, a zero-mean gaussian process is a good model for the received signal. One of the objectives of such a procedure is to estimate some of the

[1]This research was sponsored by the National Aeronautics and Space Administration under grant NsG-354.

Hard copy (HC) 2.00

Microfiche (MF) .50

ff 653 July 65

[redacted]

target's parameters (e.g., the range and rate of rotation of a planet). In this situation the total duration of observation ranges from ten to several hundred hours, yet even for the longer observation times, the quality of the estimates of some of the parameters of interest is never any more than satisfactory. Thus in this situation, both computational simplicity and asymptotic efficiency are of paramount importance, and one is led to find estimation methods having both of these qualities.

Stochastic approximation methods such as the Robbins-Munro [10] and Kiefer-Wolfowitz [8] methods, being recursive, are inherently computationally convenient. If one uses the original methods with fixed gain sequences, a_n, the fastest rate of convergence is obtained with a sequence of the form $a_n = A/n$, providing A is sufficiently large. However, the asymptotic variance of the estimates is of the form

$$A^2/n(2AG - 1)$$

in which G depends upon the unknown value of the parameter to be estimated [11]. Thus proper choice of A is quite important in obtaining the smallest possible variance, yet is difficult because of the dependence of G upon the unknown parameter value. Gardner [5] and Alberts [1][1] improve on this situation by considering nonfixed sequences of the form

$$a_n = A \quad \gamma_n / \sum_{m=1}^{n} h_n$$

for a variety of choices of the γ_n and h_n. Their work is still in flux, but at present their methods do not yield asymptotically efficient estimates when applied to the problem considered here

---

[1] A joint publication is in the process of preparation.

[redacted]

(Gardner has one method that is asymptotically efficient but whose computational complexity increases with the iteration number). For this reason, we consider here a modified Robbins-Munro method (which is essentially a stochastic Newton-Raphson method) in which the gain sequence is of the form

$$a_n = g/n$$

in which g is a function only of the previous estimate. This method requires more in the way of regularity conditions than the work of Albert and Gardner. However, for application to the problem considered here, these conditions are in general justified, and the method leads to estimates that are asymptotically efficient.

We proceed as follows. In Section 2 we present a modified Robbins-Munro procedure and prove that the mean-square error in the sequence of estimates tends to a certain limiting value. In Section 3, we return to the problem of specific interest, estimating the parameters of a covariance function, and derive explicit expressions for the partial derivatives of the log liklihood function and for the elements appearing in the Cramèr-Rao inequality. We then discuss sufficient conditions under which the method described in Section 2 is applicable and point out that, when the method is applicable, it leads to estimates that are asymptotically efficient.

2. A modified Robbins-Munro method. Let $\underset{\sim}{\alpha}$ denote an unknown M-dimensional vector parameter and let $\underset{\sim}{Y}_n(\underset{\sim}{\alpha})$, $n = 1, 2, 3, \cdots$, denote a sequence of M-dimensional vector valued random variables which are identically distributed and conditionally independent. We assume that each observation in the sequence, $\underset{\sim}{Y}_n$, can be evaluated for any choice, $\underset{\sim}{\alpha}_n$, of the parameter $\underset{\sim}{\alpha}$. Let $m_i(\underset{\sim}{\alpha})$ denote the expected value of $Y_{n,i}(\underset{\sim}{\alpha})$, the i-th component of $\underset{\sim}{Y}_n(\underset{\sim}{\alpha})$, $i = 1, 2, \cdots, M$, and

(2.1) $$\underset{\sim}{m}(\underset{\sim}{\alpha}) = E\left\{\underset{\sim}{Y}_n(\underset{\sim}{\alpha})\right\}.$$

We assume there is a value of the $\underset{\sim}{\alpha}$ parameter, say $\underset{\sim}{\theta}$, for which

(2.2) $$\underset{\sim}{m}(\underset{\sim}{\theta}) = \underset{\sim}{0}$$

and we assume further that $\underset{\sim}{\theta}$ is known to lie in the interior of some bounded rectangular subset A of R_M; i.e., $\theta_i \in (a_i, b_i)$(a_i and b_i finite). We wish to use the sequence of observations, $\underset{\sim}{Y}_n(\underset{\sim}{\alpha}_n)$, to determine $\underset{\sim}{\theta}$.

We denote by $G(\underset{\sim}{\alpha})$ a suitably chosen M by M matrix whose properties will be described later, and we consider the modified M-dimensional Robbins-Munro procedure described by the two equations

(2.3) $$\underset{\sim}{\eta}_{n+1} = \underset{\sim}{\alpha}_n + \frac{1}{n} G(\underset{\sim}{\alpha}_n)\underset{\sim}{Y}_n(\underset{\sim}{\alpha}_n)$$

and

(2.4) $$\alpha_{n+1,i} = \begin{cases} \eta_{n+1,i} & \text{if } \eta_{n+1,i} \in [a_i, b_i] \\ a_i & \text{if } \eta_{n+1,i} < a_i \\ b_i & \text{if } \eta_{n+1,i} > b_i \end{cases} \qquad i = 1, 2, \cdots, M.$$

We now state formally our assumptions on the structure of $\underset{\sim}{Y}(\underset{\sim}{\alpha})$. We denote the transpose of a matrix by a prime and use the usual notation for the Euclidean inner product and norm.

ASSUMPTION 1. The $\underset{\sim}{Y}_n(\underset{\sim}{\alpha})$ are identically distributed and conditionally independent for all n and the components of the $\underset{\sim}{Y}_n(\underset{\sim}{\alpha})$ have bounded moments of all finite order for all $\underset{\sim}{\alpha} \in A$.

ASSUMPTION 2. There exists an $\epsilon > 0$ such that

$$\theta_i \in [a_i + \epsilon, b_i - \epsilon] \qquad -\infty < a \leq a_i < b_i \leq b < \infty$$

$$i = 1, 2, \cdots, M.$$

ASSUMPTION 3.

$$\left.\frac{\partial m_i(\underset{\sim}{\alpha})}{\partial \alpha_j}\right|_{\underset{\sim}{\alpha}=\underset{\sim}{\theta}} = -E\left\{Y_{n,i}(\underset{\sim}{\theta})\, Y_{n,j}(\underset{\sim}{\theta})\right\} = -b_{ij}(\underset{\sim}{\theta}) \quad i, j = 1, 2, \cdots, M$$

the matrix $G(\underset{\sim}{\alpha})$ is symetric and has an inverse, $H(\underset{\sim}{\alpha})$, for all $\underset{\sim}{\alpha} \in A$ and

$$H(\underset{\sim}{\theta}) = G^{-1}(\underset{\sim}{\theta}) = B(\underset{\sim}{\theta}) = [b_{ij}(\underset{\sim}{\theta})] .$$

ASSUMPTION 4. There exists a K_0 and K_0', $0 < K_0 \leq K_0' < \infty$ such that

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

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

ASSUMPTION 5.

$$G(\underset{\sim}{\alpha}) \underset{\sim}{m}(\underset{\sim}{\alpha}) = -(\underset{\sim}{\alpha} - \underset{\sim}{\theta}) + \underset{\sim}{\zeta}$$

with

$$||\underset{\sim}{\zeta}|| \leq \tfrac{1}{2} K_1 \; ||\underset{\sim}{\alpha} - \underset{\sim}{\theta}||^2, \qquad K_1 < \infty$$

and

$$E\left\{\underset{\sim}{Y}_n'(\underset{\sim}{\alpha}) G'(\underset{\sim}{\alpha}) G(\underset{\sim}{\alpha}) \underset{\sim}{Y}_n(\underset{\sim}{\alpha})\right\} = \sum_{k=1}^{N} g_{kk}(\underset{\sim}{\theta}) + \tau$$

with

$$|\tau| \leq k_2 ||\underset{\sim}{\alpha} - \underset{\sim}{\theta}||, \qquad K_2 < \infty$$

(note that Assumption 3 is requisite for the two statements of this assumption to hold at $\underset{\sim}{\alpha} = \underset{\sim}{\theta}$).

The main result of this section is the following:

THEOREM 1. Assumptions 1-5 imply

$$E\left\{\|\underset{\sim}{\alpha}_n - \underset{\sim}{\theta}\|^2\right\} = \frac{1}{n}\sum_{k=1}^{M} g_{kk}(\underset{\sim}{\theta}) + 0(n^{-(1+\gamma)}), \qquad \gamma > 0.$$

We now give the proof of Theorem 1; at the conclusion of the proof we give a slight extension, Theorem 2, which concerns the convergence of the components, $\alpha_{n,i}$.

PROOF OF THEOREM 1. Subtracting $\underset{\sim}{\theta}$ from both sides of Equation (2.3), taking the norm of both sides of the resulting equation, andnoting that G is symetric, yields

$$\|\underset{\sim}{\eta}_{n+1} - \underset{\sim}{\theta}\|^2 = \|\underset{\sim}{\alpha}_n - \underset{\sim}{\theta}\|^2 + \frac{2}{n}(\underset{\sim}{\alpha}_n - \underset{\sim}{\theta})' G(\underset{\sim}{\alpha}_n)\underset{\sim}{Y}_n(\underset{\sim}{\alpha}_n) + \frac{1}{n^2} Y_n'(\underset{\sim}{\alpha}_n)G'(\underset{\sim}{\alpha}_n) G(\underset{\sim}{\alpha}_n) Y_n(\underset{\sim}{\alpha}_n). \tag{2.5}$$

Note that Assumption 2 implies that with certainty

$$|\eta_{n+1,i} - \theta_i| \geq |\alpha_{n+1,i} - \theta_i| \tag{2.6}$$

using this inequality in Equation (2.5) and taking $E\{\cdot \mid \underset{\sim}{\alpha}_n\}$ of both sides of the resulting equation gives

$$E\left\{\|\underset{\sim}{\alpha}_{n+1} - \underset{\sim}{\theta}\|^2 \mid \underset{\sim}{\alpha}_n\right\} \leq \|\underset{\sim}{\alpha}_n - \underset{\sim}{\theta}\|^2 - \frac{2}{n}(\underset{\sim}{\alpha}_n - \underset{\sim}{\theta})' G(\underset{\sim}{\alpha}_n)\underset{\sim}{m}(\underset{\sim}{\alpha}_n) + \frac{1}{n^2} E\left\{\underset{\sim}{Y}_n'(\underset{\sim}{\alpha}_n)G'(\underset{\sim}{\alpha}_n)G(\underset{\sim}{\alpha}_n)\underset{\sim}{Y}_n(\underset{\sim}{\alpha}_n) \mid \underset{\sim}{\alpha}_n\right\}. \tag{2.7}$$

Applying Assumption 5, this inequality weakens to

$$(2.8) \qquad E\left\{\|\underset{\sim}{\alpha}_{n+1} - \underset{\sim}{\theta}\|^2 \,|\, \underset{\sim}{\alpha}_n\right\} \leq \left(1 - \frac{2}{n}\right)\|\underset{\sim}{\alpha}_n - \underset{\sim}{\theta}\|^2 + (K_1/n)\|\underset{\sim}{\alpha}_n - \underset{\sim}{\theta}\|^3$$

$$+ \frac{1}{n^2} \sum_{k=1}^{M} g_{kk}(\underset{\sim}{\theta}) + (K_2/n^2)\|\underset{\sim}{\alpha}_n - \underset{\sim}{\theta}\|.$$

For brevity we will denote

$$(2.9) \qquad P = \sum_{k=1}^{M} g_{kk}(\underset{\sim}{\theta}), \qquad |X_n| = \|\underset{\sim}{\alpha}_n - \underset{\sim}{\theta}\|, \qquad b_n = E\left\{\|\underset{\sim}{\alpha}_n - \underset{\sim}{\theta}\|^2\right\}.$$

Taking expected values on both sides of inequality (2.8) and using the Schwartz inequality on the last term, we have

$$(2.10) \qquad b_{n+1} \leq b_n\left(1 - \frac{2}{n}\right) + (K_1/n)\; E\left\{|X_n|^3\right\} + \frac{P}{n^2} + (K_2/n^2)b_n^{1/2}.$$

Now the only situation of interest is the one in which

$$E\left\{ Y_n'(\underset{\sim}{\alpha})G'(\underset{\sim}{\alpha})G(\underset{\sim}{\alpha})Y_n(\underset{\sim}{\alpha}) \right\}$$

is strictly greater than zero in an $\underset{\sim}{\alpha}$ neighborhood about $\underset{\sim}{\theta}$. We will show later that under this condition b_n cannot approach zero faster than $1/n$; thus

$$(2.11) \qquad (K_2/n^2)b_n^{1/2} \leq K_3/n^{3/2}\, b_n, \qquad K_3 < \infty.$$

Now, the Hölder inequality implies

$$E\left\{|X_n|^3\right\} = E\left\{|X_n|^{1/m}\ |X_n|^{(3m-1)/m}\right\}$$

$$\leq E^{(q/q+1)}\left\{|X_n|^{(q+1)/m}\right\}\ E^{(1/q+1)}\ \left\{|X_n|^{(q+1)(3m-1)/m}\right\}$$

setting $(q+1)/qm = 2$

$$E\left\{|X_n|^3\right\} \leq E^{(q/q+1)}\left\{|X_n|^2\right\} E^{(1/q+1)}\left\{|X_n|^{q+3}\right\} \quad . \tag{2.12}$$

Substituting inequalities (2.12) and (2.11) into (2.10) yields

$$b_{n+1} \leq b_n \left\{1 - \frac{2}{n} + K_3/n^{3/2} + (K_1/n)\left[\frac{E\left\{|X_n|^{q+3}\right\}}{b_n}\right]^{(1/q+1)}\right\} + \frac{P}{n^2} \quad . \tag{2.13}$$

Our task is now to find a bound for the last term inside the brackets in inequality (2.13). Applying inequality (2.6) to Equation (2.5), raising both sides of the resulting equation to the p-th power (p an integer), taking $E\left\{\cdot\ |\underset{\sim}{\alpha}_n\right\}$ on both sides of this equation, and applying Assumption 1 yields

$$E\left\{|X_n|^{2p}|\underset{\sim}{\alpha}_n\right\} \leq |X_n|^{2p} - \frac{2p}{n}|X_n|^{2(p-1)}(\underset{\sim}{\alpha}_n - \underset{\sim}{\theta})'G(\underset{\sim}{\alpha}_n)\underset{\sim}{m}(\underset{\sim}{\alpha}_n) + \sum_{\ell=2}^{2p} C_\ell \frac{|X_n|^{2p-\ell}}{n^\ell} \tag{2.14}$$

$$C_\ell < \infty, \quad \ell = 2, 3, \cdots, 2p.$$

For convenience we will denote

$$\beta_n^p = E\left\{|X_n|^p\right\} = E\left\{||\underset{\sim}{\alpha}_n - \underset{\sim}{\theta}||^p\right\} \quad . \tag{2.15}$$

Applying Assumption 4 to inequality (2.14) and taking expected values yields

$$(2.16) \qquad \beta_{n+1}^{2p} \leq \left(1 - \frac{2pK_0}{n}\right) \beta_n^{2p} + \sum_{\ell=2}^{2p} C_\ell \, \beta_n^{2p-\ell}/n^\ell .$$

We now need a recursion inequality in the opposite direction for b_n; unfortunately, the fact that our search procedure is confined to a bounded interval will here entail an undue amount of unpleasant manipulation. Let

$$(2.17) \qquad \underset{\sim}{Z}_n(\underset{\sim}{\alpha}) = G(\underset{\sim}{\alpha}) \, \underset{\sim}{Y}_n(\underset{\sim}{\alpha})$$

and I denote the identity matrix. Then the recursion procedure can be defined by the single equation

$$(2.18) \qquad \underset{\sim}{\alpha}_{n+1} = \underset{\sim}{\alpha}_n - \frac{1}{n} \left[I - I_n(\underset{\sim}{\alpha}_n, Z_n) \right] \underset{\sim}{Z}_n(\underset{\sim}{\alpha}_n)$$

in which I_n is a diagonal matrix whose k-th entry is

$$(2.19) \qquad I_{n,k} = 0 \quad \text{if } -|\alpha_{n,k} - a_k| n \leq Z_{n,k}(\alpha_n) \leq n|b_k - \alpha_{n,k}|$$

between 0 and 1 otherwise.

Now subtract $\underset{\sim}{\theta}$ from both sides of Equation (2.18) and take the norm of both sides; if we then take $E\left\{E\left\{\cdot \mid \underset{\sim}{\alpha}_n\right\}\right\}$ and use Assumption 4, we obtain

$$(2.20) \qquad E\left\{||\underset{\sim}{\alpha}_{n+1} - \underset{\sim}{\theta}||^2\right\} \geq E\left\{||\underset{\sim}{\alpha}_n - \underset{\sim}{\theta}||^2\right\} (1 - 2K_0'/n)$$

$$- \frac{2}{n} E\left\{(\underset{\sim}{\alpha}_n - \underset{\sim}{\theta})' E\left\{ I_n(\underset{\sim}{\alpha}, \underset{\sim}{Z}_n) \underset{\sim}{Z}_n(\underset{\sim}{\alpha}_n) \mid \underset{\sim}{\alpha}_n \right\}\right\} + \frac{R}{n^2} ; R > 0.$$

We must investigate the second term in this inequality. Let

$$(2.21) \qquad J_k(\alpha_{n,k}) = 1 \quad \text{if } a_k + \epsilon/2 < \alpha_{n,k} < b_k - \epsilon/2$$

$$0 \quad \text{otherwise.}$$

Now when $J_k(\alpha_{n,k}) = 0$, we can write

$$(2.22) \qquad |E\{I_{n,k}Z_{n,k}|\underset{\sim}{\alpha}_n\}| \leq E\{|Z_{n,k}|\ |\underset{\sim}{\alpha}_n\} \leq K_{4,k}$$

in which we have used Assumption 1. Now if $J_k(\alpha_{n,k})$ is one

$$(2.23) \qquad |E\{I_{n,k}Z_{n,k}|\underset{\sim}{\alpha}_n\}| \leq \int_{|Z_{n,k}|\geq n\epsilon/2} |Z_{n,k}|dP \leq \frac{2}{n\epsilon}\int_{|Z_{n,k}|\geq n\epsilon/2} |Z_{n,k}|^2 dP$$

$$\leq K_{5,k}/n$$

and

$$(2.24) \qquad |E\{(\alpha_{n,k} - \theta_k)E\{I_{n,k}Z_{n,k}|\underset{\sim}{\alpha}_n\}\}|$$

$$\leq |E\{(\alpha_{n,k} - \theta)(1 - J_k(\alpha_{n,k}))E\{I_{n,k}Z_{n,k}|\underset{\sim}{\alpha}_n\}\}|$$

$$+ |E\{(\alpha_{n,k} - \theta_k)J_k(\alpha_{n,k})E\{I_{n,k}Z_{n,k}|\underset{\sim}{\alpha}_n\}\}|$$

$$\leq E\{|\alpha_{n,k} - \theta_k|(1 - J_k(\alpha_{n,k}))K_{4,k}\}$$

$$+ E\{|\alpha_{n,k} - \theta_k|K_{5,k}/n\}$$

$$\leq E^{1/2}\{(\alpha_{n,k} - \theta_k)^2\}[K_{4,k}E^{1/2}\{(1 - J_k(\alpha_{n,k}))^2\} + K_{5,k}/n]$$

in which we have used the Scwhartz inequality. Now using Assumption 2 and the Tchebyshev inequality

$$(2.25) \quad E\left\{(1 - J_k(\alpha_{n,k}))^2\right\} = E\left\{(1 - J_k(\alpha_{n,k}))\right\}$$

$$\leq P\left\{ |\alpha_{n,k} - \theta_k| \geq \epsilon/2 \right\} \leq \frac{4}{\epsilon^2} E\left\{(\alpha_{n,k} - \theta_k)^2\right\} .$$

Letting

$$K_4 = \max_{k=1,2,\cdots,M} \{4/\epsilon^2\} K_{4,k}, \qquad K_5 = \max_{k=1,2,\cdots,M} K_{5,k}$$

we have, combining inequalities (2.24) and (2.25)

$$(2.26) \quad \left| E \left\{ (\alpha_{n,k} - \theta_k) E\left\{ I_{n,k} \quad Z_{n,k} | \underset{\sim}{\alpha}_n \right\} \right\} \right|$$

$$\leq K_4 E\left\{ (\alpha_{n,k} - \theta_k)^2 \right\} + (K_5/n) E^{1/2}\left\{ (\alpha_{n,k} - \theta_k)^2 \right\} .$$

If we substitute inequality (2.26) into inequality (2.20), we obtain

$$(2.27) \quad b_{n+1} \geq b_n \left[1 - 2(K_0' + K_4)/n\right]$$

$$+ \frac{1}{n^2} \left[R - (2 K_5/n) \sum_{k=1}^{M} E^{1/2} \left\{ (\alpha_{n,k} - \theta_k)^2 \right\} \right].$$

But Assumptions 1 and 4 are sufficient to show via the usual arguements that

$$E\left\{ || \underset{\sim}{\alpha}_n - \underset{\sim}{\theta} ||^2 \right\} \leq K_M \; n^{-\beta}; \quad \beta > 0, \quad k_M < \infty.$$

Thus there is an N_0 such that for $n \geq N_0$ the last term in inequality (2.27) is positive and

$$(2.28) \quad b_{n+1} > b_n \left[1 - 2(K_0' + K_4)/n\right] + R/2n^2, \qquad n \geq N_0.$$

This inequality, together with Chung's second Lemma [2] implies that b_n cannot approach zero faster than $1/n$. We will use this to weaken inequality (2.26) to

$$(2.29) \qquad \left| E\left\{ (\alpha_{n,k} - \theta_k) \quad E\left\{ I_{n,k} \quad Z_{n,k} \mid \underset{\sim}{\alpha}_n \right\}\right\} \right| \leq K_6 E\left\{ (\alpha_{n,k} - \theta_k)^2 \right\}.$$

Substituting this result into inequality (2.20) and omitting the positive term, we finally obtain

$$(2.30) \qquad b_{n+1} \geq b_n \left[1 - 2(K_0' + K_6)/n\right].$$

Now

$$\left[1 - 2C/n\right]^{-1} = \left[1 + 3C/n\right] \left[1 + C/n - 6 (C/n)^2\right]^{-1}.$$

Thus, if we let N_1 be the smallest integer such that

$$6(K_0' + K_6) \leq N_1$$

then for $n \geq N_1$ we can invert inequality (2.30) to yield

$$(2.31) \quad 1/b_{n+1} \leq (1/b_n) \left[1 + 3(K_0' + K_6)/n\right] \qquad n \geq N_1.$$

Combining inequalities (2.16) and (2.31) yields for $n \geq N_1$

$$(2.32) \quad \beta_{n+1}^{2p}/b_{n+1} \leq (1 - 2pK_0/n) \left[1 + 3(K_0' + K_6)/n\right] \beta_n^{2p}/b_n$$
$$+ \left[1 + 3(K_0' + K_6)/N_1\right] \sum_{\ell=2}^{2p} C_\ell \beta_n^{2p-\ell}/b_n n^\ell .$$

Noting that $\|\underset{\sim}{\alpha}_n - \underset{\sim}{\theta}\|$ is bounded and that b_n cannot go to zero faster than $1/n$, this can be weakened to

(2.33) $$\beta^{2p}_{n+1}/b_{n+1} \le (1 - \left[2pK_0 - 3(K_0' + K_0)\right]/n)\, \beta^{2p}_n/b_n + K_7(p)/n^2$$

for $n \geq N_1$. The constants K_0, K_0', and K_6 are all independent of p; we thus let p_0 be the smallest integer such that $2p_0K_0 - 3(K_0' + K_6) \geq 2$ and set $K_7(p_0) = K_7$. Then applying Chung's first Lemma [2] to the resulting recursive inequality

(2.34) $$\beta_n^{2p_0}/b_n \le K_7/n + 0(1/n^2) \le K_8/n \quad \text{for} \quad n \ge N_1 .$$

We now return to inequality (2.13) and set $q + 3 = 2p_0$ and $\gamma = \left[2(p_0 - 1)\right]^{-1}$, where $0 < \gamma \leq 1/2$ since $2p_0 \geq 4$. Then substituting inequality (2.3.4) into (2.13)

(2.35) $$b_{n+1} \le b_n \left[1 - 2/n + K_9/n^{1+\gamma}\right] + P/n^2$$

$$n \ge N_1 \qquad K_9 < \infty .$$

Now for a recursion relation of the form

(2.36) $$x_{k+1} \le (1 - a_k)\, x_k + \epsilon_k$$

it is well known that if $1 - a_k$ is positive for all $k \geq N$, then

(2.37) $$x_n \le x_N\, \beta_{N-1,\, n-1} + \sum_{k=n}^{n-1} \epsilon_k \beta_{k,\, N-1}$$

in which

(2.38) $$\beta_{mn} = \begin{cases} \prod_{j=m+1}^{n} (1 - a_j) & N \le m < n \\ 1 & m = n \end{cases}$$

Using the convexity of log x we have for $0 \leq a_j \leq 1$ that $\log(1 - a_j) \leq -a_j$, thus

$$(2.39) \qquad \beta_{mn} \leq \exp\left\{ - \sum_{j=m+1}^{n} a_j \right\} .$$

Let N_2 be the smallest integer such that $0 \leq + 2/n - K_9/n^{\gamma+1} \leq 1$ for all $n \geq N_2$ and let $N = \max \; N_1, N_2$. For the recursive inequality (2.35) we thus have

$$(2.40) \quad \beta_{m,n-1} \leq \exp\left\{ - \sum_{K=m+1}^{n-1} (2/k - K_9/k^{\gamma+1}) \right\}; n, m \geq N .$$

Bounding the sum by appropriate integrals and setting $C = K_9/\gamma$, yields

$$(2.41) \quad \beta_{m,n-1} \leq (m + 1/n)^2 \left\{ \exp \; C/m^{\gamma} \right\} \qquad m, n \geq N .$$

Substituting the appropriate quantities from inequalities (2.35) and (2.41) into inequality (2.37) yields

$$(2.42) \quad b_n \leq b_N (N/n)^2 \exp(C/N^{\gamma}) + (P/n^2) \sum_{k=N}^{n-1} (1 + 2/k + 1/k^2) \exp(C/k^{\gamma}) \qquad n \geq N .$$

The sum appearing in this inequality can be bounded by

$$(2.43) \qquad \Sigma \leq n - N + \int_{N-1}^{n-1} \left[\exp(C/t^{\gamma}) - 1 \right] dt + \exp\left[C/(N-1)^{\gamma} \right] \int_{N-1}^{n-1} (2/t + 1/t^2)\, dt .$$

In the range of intereset for the integral, we can use the bound

$$\exp\left\{C/t^{\gamma}\right\} -1 \le t^{-\gamma} \quad (\exp\left[C/(N-1)^{\gamma}\right] -1)\ (N-1)^{\gamma}/C \quad .$$

Using this bound to evaluate an upper bound for the first integral in inequality (2.43), and evaluating the second integral directly, we can substitute the result into inequality (2.42) to yield finally

$$(2.44) \qquad b_n \le P/n + K_{10}/n^{1+\gamma} \qquad n \ge N \quad .$$

This completes the proof of Theorem 1.

We now seek bounds on the individual terms $E\left\{(\alpha_{n,k} - \theta_k)^2\right\}$, $k = 1, 2, \ldots, M$. We could obtain such a bound for the $k = q$ term by subtracting the Cramer-Rao bounds for the terms $k \neq q$ from the right-hand side of the bound of Theorem 1. However, since the $\alpha_{n,k}$ are not unbiased, these Cramer-Rao bounds would require the evaluation of the partials of the bias terms. This is an unpleasant prospect, and we consider another approach.

Consider defining the unknown $\underset{\sim}{\alpha}$ parameters by the new set of variables

$$(2.45) \qquad \underset{\sim}{\beta} = C\,\underset{\sim}{\alpha}$$

in which C is a non-singular matrix. In terms of the $\underset{\sim}{\beta}$ variables, the pertinent quantities are

$$(2.46) \qquad \begin{aligned} H(\underset{\sim}{\alpha}) &= C'\,H(\underset{\sim}{\beta})\,C, \quad G(\underset{\sim}{\alpha}) = C^{-1}\,G(\underset{\sim}{\beta})\,(C')^{-1} \\ \underset{\sim}{Y}(\underset{\sim}{\alpha}) &= C'\,\underset{\sim}{Y}(\underset{\sim}{\beta}) \quad \text{and} \quad \underset{\sim}{m}(\underset{\sim}{\alpha}) = C'\,\underset{\sim}{m}(\underset{\sim}{\beta}) \end{aligned}$$

in which ' denotes transpose. These expressions follow directly from the definitions of the quantities involved and the relation

$$\partial/\partial\alpha_i = \sum_{j=1}^{M} c_{ji}\ \partial/\partial\beta_j \qquad i = 1, 2, \ldots, M\ .$$

Using the relations (2.45) and (2.46) one can verify directly that the first recursion relation for the $\underset{\sim}{\beta}_n$ is identical in form to Eq. (2.3). Also, if (and only if) C is diagonal, then the second recursion relation for the $\underset{\sim}{\beta}_n$ is identical in form to Eq. (2.4). Direct evaluation using Eqs. (2.45) and (2.46) also shows that Assumptions (1), (2), (3), and (5) hold either directly or in suitably equivalent form for the $\underset{\sim}{\beta}$ variables when C is symetric (and hence in particular when C is diagonal). Regarding Assumption (4), let C be diagonal and let $\underset{\sim}{\theta}_\beta = C\underset{\sim}{\theta}$. Then

$$(2.47)\quad (\underset{\sim}{\beta} - \underset{\sim}{\theta}_\beta)' \, G(\underset{\sim}{\beta}) \, \underset{\approx}{m}(\underset{\sim}{\beta}) = (\underset{\sim}{\alpha} - \underset{\sim}{\theta})' \, C \, C \, G(\underset{\sim}{\alpha}) \, C \, C^{-1} \, \underset{\approx}{m}(\underset{\sim}{\alpha})$$

$$= \sum_{k=1}^{M} (c_{kk})^2 (\alpha_k - \theta_k) \left[G(\underset{\sim}{\alpha}) \, \underset{\approx}{m}(\underset{\sim}{\alpha})\right]_k .$$

But by Assumption (4) each individual term in this sum is non-negative, thus

$$(2.47)\quad C^2_{min}(\underset{\sim}{\alpha} - \underset{\sim}{\theta})' \, G(\underset{\sim}{\alpha}) \, \underset{\approx}{m}(\underset{\sim}{\alpha}) \le (\underset{\sim}{\beta} - \underset{\sim}{\theta}_\beta)' G(\underset{\sim}{\beta}) \, \underset{\approx}{m}(\underset{\sim}{\beta})$$

$$\le C^2_{max}(\underset{\sim}{\alpha} - \underset{\sim}{\theta})' \, G(\underset{\sim}{\alpha}) \, \underset{\approx}{m}(\underset{\sim}{\alpha})$$

in which

$$C^2_{min} = \min_{k=1,2,..,M} c^2_{kk} \; ; \quad C^2_{max} = \max_{k=1,2,..,M} c^2_{kk} .$$

Similarly

$$(2.48)\quad (1/C^2_{max}) \, || \underset{\sim}{\beta} - \underset{\sim}{\theta}_\beta ||^2 \le || \underset{\sim}{\alpha} - \underset{\sim}{\theta} ||^2 \le (1/C^2_{min}) \, || \underset{\sim}{\beta} - \underset{\sim}{\theta}_\beta ||^2 .$$

Combining inequalities (2.47) and (2.48) shows that Assumption (4) again holds in the $\underset{\sim}{\beta}$ coordinates with K_0 replaced by $K_0\,(C^2_{min}/C^2_{max})$ and K'_0 replaced by $K'_0\,(C^2_{max}/C^2_{min})$.

We can now state

THEOREM 2. Assumptions (1) - (5) imply

$$E\left\{(\alpha_{n,i} - \theta_i)^2\right\} \leq g_{ii}(\underset{\sim}{\theta})/n + \left(0(n^{-(1+\gamma)}\right), \quad \gamma > 1 \quad i = 1, 2, \ldots, M.$$

PROOF OF THEOREM 2. From the above discussion, it follows that Theorem 1 still holds for any nonsingular diagonal transformation of the original coordinates; thus

$$(2.49) \qquad \sum_{i=1}^{M} c_i^2 \; E\left\{(\alpha_{n,}\, i - \theta_i)^2\right\} \leq (1/n) \sum_{i=1}^{M} c_i^2\, g_{ii}(\theta) + K(C)\,(1/n^{1+\gamma}).$$

Now pick M sets of the c_i with

$$(2.50) \qquad \begin{array}{lll} c_1^j = 1 & j = 1, 2, \ldots, M & \\ c_i^j = 1 & i \neq j & \\ \quad\; 1 + \epsilon_0 & i = j & 0 < \epsilon_0 < 1 \\ & i = 2, 3, \ldots, M; & j = 1, 2, \ldots, M . \end{array}$$

Then the relations of inequality (2.49) for these M choices of the C matrix are expressed

$$(2.51) \qquad \sum_{i=1}^{M} (c_i^j)^2 \left[E\left\{(\alpha_{n,i} - \theta_i)^2\right\} - g_{ii}(\underset{\sim}{\theta})/n\right] = f(C_j)$$

$$j = 1, 2, \ldots, M$$

with

$$0 \leq f(C_j) \leq K(C_j)/n^{1+\gamma} \quad .$$

But the $M \times M$ matrix of c_i^j appearing in the set of M linear Equations (2.51) is nonsingular; this set may be solved to yield

$$E\left\{(\alpha_{n,i} - \theta_i)^2\right\} = g_{ii}(\underset{\sim}{\theta})/n + \sum_{j=1}^{M} q_{ij}\, f(C_j)$$

in which the q_i's are all bounded. Thus Theorem 2.

3. The log-likelihood function and the Cramer-Rao inequality for the covariance problem. We turn in this section to the problem of direct interest; estimation of the parameters of a covariance function. Our goal is two-fold: to obtain expressions for the partial derivatives of the covariance function and the matrix appearing in the Cramer-Rao bound, and to determine conditions which are sufficient to guranatee that the theorems of Section 2 are applicable here.

We begin by introducing the model that we consider and the necessary notation. We assume that our observed process $W(t)$ is the sum of the information bearing signal $S(t)$ which depends on the parameters $\alpha_1, \alpha_2, \ldots, \alpha_M$, plus an interfering "white" gaussian noise process $N(t)$, specifically.

Condition (i). $W(t) = S(t) + N(t)$ in which S and N are independent zero-mean gaussian processes. The process N is "white" in the sense that for any square integrable $f(t)$ and $g(t)$

$$E\left\{\int_0^T \int_0^T N(t)\, N(s)\, g(t)\, f(s)\, dt\, ds\right\} = N \int_0^T g(t)\, f(t)\, dt$$

and, for all square integrable $h(s,t)$

$$E\left\{\underbrace{\int_0^T \cdots \int_0^T}_{2N \text{ times}} \prod_{i=1}^{N} N(t_i)\, N(s_i)\, h(t_i, s_i)\, dt_1 \ldots dt_N\, ds_1 \ldots ds_N\right\}$$

$$= \sum \underbrace{\int_0^T \cdots \int_0^T}_{N \text{ times}} \prod_{i=1}^{N} h(\cdot\,,\,\cdot)\, dt_1 \ldots dt_N$$

in which the sum is over all ways in which N variables can occur as 2N arguments.

As before, we let $\underset{\sim}{\alpha}$ be the M dimensional vector whose components represent the possible values of the parameters $\alpha_1, \alpha_2, \ldots, \alpha_M$, and $\underset{\sim}{\theta}$ denote the value of $\underset{\sim}{\alpha}$ that actually pertains to the process being observed. The process W(t) is observed on the disjoint time intervals $I_1, I_2, I_3, \ldots$, each of duration T. We denote the functional dependence of the covariance function of S(t) on $\underset{\sim}{\alpha}$ by $\phi_s(t, s, \underset{\sim}{\alpha})$, $t, s \epsilon [0, t]$. Thus

$$(3.1) \qquad E\{W_t W_s\} = \phi_s(t - \tau_n, s - \tau_n, \underset{\sim}{\theta}) + N\,\delta(t - s)$$

$$t, s, \epsilon\ I_n; \qquad \tau_n \text{ the initial point of } I_n;\ n = 1, 2, 3, \ldots$$

in which the dirac-delta function is to be regarded as a distribution on the space of continuous functions.

Let $\psi_n(t, \underset{\sim}{\alpha})$ and $\lambda_n(\underset{\sim}{\alpha})$ be the Karhunen-Loéve expansion corresponding to $\phi_s(t, s, \underset{\sim}{\alpha})$; i.e., the $\psi_n(t, \underset{\sim}{\alpha})$ and $\lambda_n(\underset{\sim}{\alpha})$ are the normalized eigenfunctions and eigenvalues of the integral equation

$$(3.2) \qquad \lambda_n(\underset{\sim}{\alpha})\,\psi_n(r, \underset{\sim}{\alpha}) = \int_0^T \phi_s(r, s, \underset{\sim}{\alpha})\,\psi_n(s, \underset{\sim}{\alpha})\,ds \qquad \begin{array}{l} 0 < r < T \\ n = 1, 2\ldots \end{array}$$

We take the likelihood function of $\underset{\sim}{\alpha}$ based on an observation W(t), $t \epsilon [0, T]$, to be the ratio of the "probability density for W(t)" under the hypothesis that the observation was at level $\underset{\sim}{\alpha}$ to the "probability density for W(t)" under the hypothesis that W(t) was noise alone. It is shown [6,9] that the natural log of this likelihood function is given by

$$(3.3) \qquad \ell(\underset{\sim}{\alpha}) = (1/2N) \sum_{n=1}^{\infty} W_n^2 \left[\lambda_n(\underset{\sim}{\alpha})/\lambda_n(\underset{\sim}{\alpha}) + N\right]$$

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

in which

$$(3.4) \qquad W_j = \int_0^T \psi_n(t, \underset{\sim}{\alpha})\, W(t)\, dt$$

(we will shortly make an assumption sufficient to guarantee the convergence in the mean square of the sum in Equation (3.3)). If we define the function

$$(3.5) \quad h(s, t, \underset{\sim}{\alpha}) = \sum_{n=1}^{\infty} \left[\lambda_n(\underset{\sim}{\alpha})/\lambda_n(\underset{\sim}{\alpha}) + N\right] \psi_n(t, \underset{\sim}{\alpha})\, \psi_n(s, \underset{\sim}{\alpha})$$

which is the solution of the integral equation

$$(3.6) \quad N h(r, t, \underset{\sim}{\alpha}) + \int_0^T \phi_s(r, s, \underset{\sim}{\alpha})\, h(s, t, \underset{\sim}{\alpha})\, ds = \phi_s(r, t, \underset{\sim}{\alpha}); r, t \in [0, T] \; .$$

Then $\ell(\underset{\sim}{\alpha})$ can be written

$$(3.7) \quad \ell(\underset{\sim}{\alpha}) = (1/2N) \int_0^T \int_0^T h(s, t, \underset{\sim}{\alpha}) W(t) W(s)\, dt\, ds - 1/2 \ln\left[\prod_{n=1}^{\infty} \left[T_n(\underset{\sim}{\alpha}) + N/N\right]\right] \; .$$

We will use the following convenient notation for certain partial derivatives

$$(3.8) \quad \phi_s^i(t, s, \underset{\sim}{\alpha}) = \partial/\partial\alpha_i\, \phi_s(t, s, \underset{\sim}{\alpha}) \qquad h^i(t, s, \underset{\sim}{\alpha}) = \partial/\partial\alpha_i\, h(t, s, \underset{\sim}{\alpha})$$

$$h^{ij}(t, s, \underset{\sim}{\alpha}) = \partial^2/\partial\alpha_i\, \partial\alpha_j \left[h(t, s, \underset{\sim}{\alpha})\right] \; .$$

Our attention will be confined to situations satisfying the following conditions:

Condition (ii). $\underset{\sim}{\theta}$ is known a-priori to lie in the interior of a bounded rectangle A (in R_M).

Condition (iii). Samples of $W(t)$ on each of the intervals $I_1, I_2, I_3, \ldots$ are all identically distributed and samples from distinct intervals are statistically independent.

Condition (iv). For all $\underset{\sim}{\alpha} \epsilon A$, the functions $h(t, s, \underset{\sim}{\alpha})$, $h^i(t, s, \underset{\sim}{\alpha})$, and $\phi_s(t, s, \underset{\sim}{\alpha})$ are continuous in t and s on $[0, T] \times [0, T]$ and the functions $\phi^i(t, s, \underset{\sim}{\alpha})$ and $h^{ij}(t, s, \underset{\sim}{\alpha})$ are L_2 on $[0, T] \times [0, T]$ for $i, j = 1, 2, \ldots, M$.

Having stated the above conditions, we can proceed with obtaining concise expressions for the partial derivatives of $\ell(\underset{\sim}{\alpha})$ and the second moments of these quantities. These will be of interest in their own right and in showing that $\ell(\underset{\sim}{\alpha})$ possess certain regularity properties. Following this, we describe a set of variables which will relate our problem at hand to the recursive estimation method of the second section. Lastly, we state the remainder of the conditions required to guarantee that the statements of Theorems 1 and 2 apply to the present situation.

We start somewhat obliquely by multiplying both sides of Equation (3.5) by $\psi_m(s, \underset{\sim}{\alpha})$, integrating both sides of the resulting equation with respect to s over $[0, T]$, and noting that (since the $\psi_n(s, \underset{\sim}{\alpha})$ are L_2 on $0, T$) we can interchange integration and summation; thus

$$(3.09) \quad \int_0^T h(s, t, \underset{\sim}{\alpha}) \, \psi_m(s, \underset{\sim}{\alpha}) \, ds = \frac{\lambda_m(\underset{\sim}{\alpha})}{\lambda_m(\underset{\sim}{\alpha}) + N} \, \psi_m(t, \underset{\sim}{\alpha}) \qquad t \epsilon [0, T] .$$

Let us now take the partial of both sides of this equation with respect to α_i and assume for the moment that the derivative of the integrand exists and that interchange of integration and differentiation is justified. This yields, upon rearrangement of terms

$$(3.10)\quad \frac{\partial}{\partial\alpha_i}\left(\frac{\lambda_m(\underset{\sim}{\alpha})}{\lambda_m(\underset{\sim}{\alpha}) + N}\right)\psi_m(t, \underset{\sim}{\alpha}) - \int_0^T h^i(r, t, \underset{\sim}{\alpha})\,\psi_m(r, \underset{\sim}{\alpha})\,dr$$

$$= -\frac{\lambda_m(\underset{\sim}{\alpha})}{\lambda_m(\underset{\sim}{\alpha}) + N}\,\psi_m^i(t, \underset{\sim}{\alpha}) + \int_0^T h(r, t, \underset{\sim}{\alpha})\,\psi_m^i(r, \underset{\sim}{\alpha})\,dr.$$

If we regard this as an integral equation in the unknown function $\psi_m^i(t, \underset{\sim}{\alpha})$, the properties of $h(r, t, \underset{\sim}{\alpha})$ assumed in Condition (iv) guarantee that $\psi_m^i(t, \underset{\sim}{\alpha})$ is L_2 on $[0, T]$ for all $\underset{\sim}{\alpha} \epsilon A$. Thus, since ψ_m, h, h^i, and ψ_m^i are all L_2 on $[0, T] \times [0, T]$ for all $\underset{\sim}{\alpha} \epsilon A$, the function

$$\frac{\partial}{\partial\alpha_i}\left[h(s, t, \underset{\sim}{\alpha})\,\psi_m(s, \underset{\sim}{\alpha})\right]$$

is L_1 on $[0, T] \times [0, T]$ and the interchange of differentiation and integration was justified [3]. Now multiply both sides of Equation (3.10) by $\psi_m(t, \underset{\sim}{\alpha})$ and integrate with respect to t over $[0, T]$. Noting by Condition (iv) and our comments above that $h(r, t)\,\psi_m^i(r, \underset{\sim}{\alpha})\,\psi_m(t, \underset{\sim}{\alpha})$ is L_1 on $[0, T] \times [0, T]$, we can apply Fubini's Theorem to the second term on the right-hand side of the equation and interchange order of integration. By Equation (3.09) the two terms on the right-hand side of the resulting equation cancel and

$$(3.11)\quad \frac{\partial}{\partial\alpha_i}\,\frac{\lambda_m(\underset{\sim}{\alpha})}{\lambda_m(\underset{\sim}{\alpha}) + N} = \int_0^T\int_0^T h^i(r, t, \underset{\sim}{\alpha})\,\psi_m(r, \underset{\sim}{\alpha})\,\psi_m(t, \underset{\sim}{\alpha})\,dt\,dr\ .$$

Carrying out the indicated differentiation, multiplying both sides of the resulting equation by $\lambda_m(\underset{\sim}{\alpha}) + N/N$, and summing yields

$$(3.12) \quad \sum_{m=1}^{\infty} \frac{\partial \lambda_m(\underset{\sim}{\alpha})}{\partial \alpha_i} \Big/ \left[\lambda_m(\underset{\sim}{\alpha}) + N\right] = \frac{1}{N} \sum_{m=1}^{\infty} \lambda_m(\underset{\sim}{\alpha}) + N$$

$$\int_0^T \int_0^T h^i(r, t, \underset{\sim}{\alpha})\, \psi_m(r, \underset{\sim}{\alpha})\, \psi_m(t, \underset{\sim}{\alpha})\, dt\, dr \quad .$$

The function $h^i(r, t, \underset{\sim}{\alpha})$ is assumed continuous in r and t for all $\underset{\sim}{\alpha} \epsilon A$; thus

$$(3.13) \quad \sum_{m=1}^{\infty} \frac{\partial \lambda_m(\underset{\sim}{\alpha})}{\partial \alpha_i} \Big/ \left[\lambda_m(\underset{\sim}{\alpha}) + N\right] = \frac{1}{N} \int_0^T \int_0^T h^i(r, t, \underset{\sim}{\alpha})$$

$$\left[N\ \delta(t - r) + \phi_s(t, r, \underset{\sim}{\alpha})\right] dt\ dr = \frac{1}{N} \int_0^T \int_0^T h^i(r, t, \underset{\sim}{\alpha})\, \phi_w(t, r, \underset{\sim}{\alpha}) dt\, dr \ .$$

Now return to Equation (3. 7) and take the partial derivative with respect to α_i:

$$(3.14) \quad \frac{\partial \ell(\underset{\sim}{\alpha})}{\partial \alpha_i} = \frac{1}{2N} \int_0^T \int_0^T h^i(s, t, \underset{\sim}{\alpha})\ W(t)\ W(s) - \frac{1}{2} \sum_{n=1}^{\infty}$$

$$\frac{\partial \lambda_n(\underset{\sim}{\alpha})}{\partial \alpha_i} \Big/ \left[\lambda_n(\underset{\sim}{\alpha}) + N\right] \quad .$$

We have interchanged differentiation and integration since, by the continuity of $h^i(s, t, \underset{\sim}{\alpha})$ and the structure of W imposed by Condition (i), the random variable

$$\frac{\partial}{\partial \alpha_i} \int_0^T \int_0^T W(t)\, W(s)\, h(s,t,\underset{\sim}{\alpha})\, dt\, ds \;-\; \int_0^T \int_0^T W(t)\, W(s)\, h^i(s,t,\underset{\sim}{\alpha})\, dt\, ds$$

has all moments of finite order equal to zero. The sum in Equation (3.14) converges by virtue of Equation (3.13). Combining Equations (3.13) and (3.14), we have finally

$$(3.15) \quad \frac{\partial \ell(\underset{\sim}{\alpha})}{\partial \alpha_i} = \frac{1}{2N} \int_0^T \int_0^T h^i(s,t,\underset{\sim}{\alpha}) \left[W(t)\, W(s) - \phi_W(t,s,\underset{\sim}{\alpha}) \right] dt\, ds \, .$$

We now wish to evaluate the matrix appearing in the Cramer-Rao bound. Let

$$(3.16) \quad B(\underset{\sim}{\alpha}) = \left[b_{ij}(\underset{\sim}{\alpha}) \right] \; ; \quad b_{ij}(\underset{\sim}{\alpha}) = E\left\{ \frac{\partial \ell(\underset{\sim}{\alpha})}{\partial \alpha_i} \; \frac{\partial \ell(\underset{\sim}{\alpha})}{\partial \alpha_j} \right\} .$$

Using Equation (3.15) and Condition (i), we have

$$(3.17) \quad b_{ij}(\underset{\sim}{\alpha}) = \frac{1}{4N^2} \iint_0^T\!\!\iint h^i(t,s,\underset{\sim}{\alpha})\, h^j(u,v,\underset{\sim}{\alpha})$$

$$\Big[\phi_W(t,u,\underset{\sim}{\theta})\, \phi_W(s,v,\underset{\sim}{\theta}) + \phi_W(t,v,\underset{\sim}{\theta})\, \phi_W(s,u,\underset{\sim}{\theta})\, \phi_W(t,s,\underset{\sim}{\theta})\, \phi_W(u,v,\underset{\sim}{\theta})$$

$$- \phi_W(t,s,\underset{\sim}{\alpha})\, \phi_W(u,v,\underset{\sim}{\alpha}) \Big] \; dt\, ds\, du\, dv$$

in which we have used ϕ_W symbolically. Using the symmetry of the functions involved, this reduces, for $\underset{\sim}{\alpha} = \underset{\sim}{\theta}$, to

$$(3.18) \quad b_{ij}(\underset{\sim}{\theta}) = \frac{1}{2N^2} \iint_0^T\!\!\iint h^i(t,s,\underset{\sim}{\theta})\, h^j(u,v,\underset{\sim}{\theta})\, \phi_W(t,u,\underset{\sim}{\theta})\, \phi_W(s,v,\underset{\sim}{\theta})\, dt\, ds\, du\, dv.$$

We seek a simpler form for this expression. To minimize the manipulation we use the (symbolic) form

$$\phi_w(t, s, \underset{\sim}{\alpha}) = N\,\delta(t - s) + \phi_s(t, s, \underset{\sim}{\alpha})$$

and will formally interchange operations. These interchanges could be justified by lengthier operations, for example, we rewrite Equation (3.6) as

$$\text{(3.19)} \qquad \int_0^T \phi_w(r, s, \underset{\sim}{\alpha})\, h(s, t, \underset{\sim}{\alpha})\, ds = \phi_s(r, t, \underset{\sim}{\alpha})$$

multiply both sides of this equation by $h^i(r, q, \underset{\sim}{\theta})\,\phi_w(q, t, \underset{\sim}{\theta})$, take the partial $\partial/\partial\alpha_j$, and evaluate at $\underset{\sim}{\alpha} = \underset{\sim}{\theta}$. This yields

$$\text{(3.20)} \qquad h^i(r, q, \underset{\sim}{\theta})\,\phi_w(q, t, \underset{\sim}{\theta}) \int_0^T h^j(s, t, \underset{\sim}{\theta})\,\phi_w(r, s, \underset{\sim}{\theta})\, ds$$

$$+\; h^i(r, q, \underset{\sim}{\theta})\,\phi_w(q, t, \underset{\sim}{\theta}) \int_0^T h(s, t, \underset{\sim}{\theta})\,\phi_s^j(r, s, \underset{\sim}{\theta})\, ds$$

$$= \phi_s^j(r, t, \underset{\sim}{\theta})\, h^i(r, q, \underset{\sim}{\theta})\,\phi_w(q, t, \underset{\sim}{\theta}) \;.$$

Now integrating both sides with respect to $r, q,$ and t, interchanging order of integration in the second term, using Equation (3.19)to cancel terms and comparing with Equation (3.18) yields finally

$$\text{(3.21)} \qquad b_{ij}(\underset{\sim}{\theta}) = \frac{1}{2N} \int_0^T \int_0^T h^i(r, t, \underset{\sim}{\theta})\,\phi_s^j(r, t, \underset{\sim}{\theta})\, dr\, dt \;.$$

Now let us make the definition

$$\text{(3.22)} \qquad m_i(\underset{\sim}{\alpha}) = E\left\{ \frac{\partial \ell(\underset{\sim}{\alpha})}{\partial \alpha_i} \right\} \;.$$

From Equation (3.15) and Condition (i) we have

$$(3.23)\quad m_i(\underset{\sim}{\alpha}) = \frac{1}{2N} \int_0^T\!\!\int h^i(s,t,\underset{\sim}{\alpha})\left[\phi_w(t,s,\underset{\sim}{\theta}) - \phi_w(t,s,\underset{\sim}{\alpha})\right] dt\, ds$$

$$= \frac{1}{2N} \int_0^T\!\!\int h^i(s,t,\underset{\sim}{\alpha})\left[\phi_s(t,s,\underset{\sim}{\theta}) - \phi_s(t,s,\underset{\sim}{\alpha})\right] dt\, ds\ .$$

Let us now take the partial derivative of both sides of this equation, noting that Condition (iv) justifies the interchange of operations [3].

$$(3.24)\quad \frac{\partial}{\partial\alpha_j} m_i(\underset{\sim}{\alpha}) = \frac{1}{2N} \int_0^T\!\!\int h^{ij}(s,t,\underset{\sim}{\alpha})\left[\phi_s(t,s,\underset{\sim}{\theta}) - \phi_s(t,s,\underset{\sim}{\alpha})\right] dt\, ds$$

$$- \frac{1}{2N} \int_0^T\!\!\int h^i(s,t,\underset{\sim}{\alpha})\left[\phi_s^j(t,s,\underset{\sim}{\alpha})\right] dt\, ds$$

and hence

$$(3.25)\quad -\frac{\partial}{\partial\alpha_j} m_i(\underset{\sim}{\alpha})\Bigg|_{\underset{\sim}{\alpha}=\underset{\sim}{\theta}} = \frac{1}{2N} \int_0^T\!\!\int h^i(s,t,\underset{\sim}{\theta})\,\phi_s^j(s,t,\underset{\sim}{\theta})\, dt\, ds$$

$$= b_{ij}(\underset{\sim}{\theta})\ .$$

Equations (3.15) and (3.25) constitute the principal results of this section. The equality of $\partial/\partial\alpha_j\, m_i$ and $-b_{ij}$ will be of interest in considering the recursive estimation procedure. The expressions for the $b_{ij}(\underset{\sim}{\theta})$ are of interest in that they appear in the Cramer-Rao inequality: if $\hat{\underset{\sim}{\theta}}$ is any unbiased estimate of $\underset{\sim}{\theta}$ based on $W(t)$, $t \epsilon\ 0, T$, then

$$(3.26)\quad \underset{\sim}{c}'\, E\left\{(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})'\right\}\underset{\sim}{c} > \underset{\sim}{c}'\, B^{-1}(\underset{\sim}{\theta})\, \underset{\sim}{c}$$

in which prime denotes transpose and $\underset{\sim}{c}$ is an arbitrary column vector. Our expressions above have been obtained for an observation on an arbitrary time interval. If the time interval in these expressions is taken to be a composite of n subintervals, each of length T; $W(t)$ on one subinterval is taken to be statistically independent of $W(t)$ on the other subintervals; and $\underset{\sim}{\theta}_n$ taken to represent an arbitrary unbiased estimate of $\underset{\sim}{\theta}$ based on $W(t)$ on the composite interval, then Equation (3. 26) becomes

$$(3.27) \quad \underset{\sim}{c}' \, E\left\{(\hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta})(\hat{\underset{\sim}{\theta}}_n - \underset{\sim}{\theta})'\right\} \underset{\sim}{c} \geq \frac{1}{n} \underset{\sim}{c}' \, B^{-1}(\underset{\sim}{\theta}) \, \underset{\sim}{c}$$

in which the elements of $B(\underset{\sim}{\theta})$ are given by Equation (3. 25), the integration being over a single subinterval of length T.

We now turn to applying the recursive method of Section 2 to the problem of interest. We will take $\underset{\sim}{Y}_n(\underset{\sim}{\alpha})$ to be the M-dimensional vector-valued observation whose i-th component is $\partial \ell / \partial \alpha_i$ evaluated on the interval I_n at the level $\underset{\sim}{\alpha}$; i. e.,

$$(3.28) \quad Y_{n,i}(\underset{\sim}{\alpha}) = \frac{1}{2N} \int_0^T \int_0^T h^i(t, s, \underset{\sim}{\alpha}) \Big[W(t + \tau_n) \, W(s + \tau_n)$$

$$- \, \phi_w(t, s, \underset{\sim}{\alpha}) \Big] \, dt \, ds \quad .$$

As in Section 2, we denote

$$\underset{\approx}{m}(\underset{\sim}{\alpha}) = E\left\{\underset{\sim}{Y}_n(\underset{\sim}{\alpha})\right\} \quad .$$

Note that by Equation (3. 23) $\underset{\approx}{m}(\underset{\sim}{\theta}) = \underset{\sim}{0}$.

Let $H(\underset{\sim}{\alpha})$ be the $M \times M$ matrix whose i-jth element is

$$(3.29) \quad h_{ij}(\underset{\sim}{\alpha}) = \frac{1}{2N} \int_0^T \int_0^T h^i(t, s, \underset{\sim}{\alpha}) \, \phi_s^j(t, s, \underset{\sim}{\alpha}) \, dt \, ds$$

and let $g_{ij}(\underset{\sim}{\alpha})$ denote the elements of $G(\underset{\sim}{\alpha}) = H^{-1}(\underset{\sim}{\alpha})$. This matrix G will be the G matrix of Section 2. Note that, comparing Equations (3. 25) and (3. 29)

$$G(\underset{\sim}{\theta}) = B^{-1}(\underset{\sim}{\theta}) \quad . \tag{3.30}$$

Thus, when the assumptions of Theorem 1 are satisfied for the above choice of $H(\underset{\sim}{\alpha})$ and $\underset{\sim}{Y}_n(\underset{\sim}{\alpha})$, Theorem 2 and inequality (3. 27) show that the estimates, $\alpha_{n,i}$, generated by the method of Section 2 have a mean square error equal to that of the "best possible" unbiased estimate.

Three remaining conditions are sufficient to guarantee that the $\underset{\sim}{Y}_n(\underset{\sim}{\alpha})$ and $G(\underset{\sim}{\alpha})$ defined above satisfy the assumptions of Section 2 and thus guarantee that the statements of Theorems 1 and 2 are applicable here. These are,

Condition (v). $H(\underset{\sim}{\alpha})$ is invertible and the elements $g_{ij}(\underset{\sim}{\alpha})$, of the inverse are uniformly bounded for all $\underset{\sim}{\alpha} \epsilon A$.

Condition (vi). For all $\underset{\sim}{\alpha} \epsilon A$

$$\left| \frac{\partial^2}{\partial\alpha_k \partial\alpha_\ell} \sum_{j=1}^{M} g_{ij}(\underset{\sim}{\alpha})\, m_j(\underset{\sim}{\alpha}) \right| \leq \frac{K_1}{\sqrt{M}} < \infty$$

$$\text{for } i, k, \ell = 1, 2, \ldots, M ,$$

and

$$\left| \frac{\partial}{\partial\alpha_\ell} \sum_{i,j,k=1}^{M} b_{ik}(\underset{\sim}{\alpha})\, g_{ji}(\underset{\sim}{\alpha})\, g_{jk}(\underset{\sim}{\alpha}) \right| \leq \frac{K_2}{\sqrt{M}} < \infty$$

$$\text{for } \ell = 1, 2, \ldots, M .$$

Condition (vii). We assume that Assumption (iv) of Section 2 holds directly for $G(\underset{\sim}{\alpha})$ and $\underset{\sim}{m}(\underset{\sim}{\alpha})$ as defined here.

Condition (vi) could be reudced to a more basic condition on the higher-order partial derivatives of $h(t, s, \underset{\sim}{\alpha})$ and $\phi_s(t, s, \underset{\sim}{\alpha})$; however, this seems of dubious value. Regarding Condition (vii), note that by Equations (3.25) and (3.30)

$$\lim_{\underset{\sim}{\alpha} \to \underset{\sim}{\theta}} (\underset{\sim}{\alpha} - \underset{\sim}{\theta})' \, G(\underset{\sim}{\alpha}) \, \underset{\approx}{m}(\underset{\sim}{\alpha}) = \lim_{\underset{\sim}{\alpha} \to \underset{\sim}{\theta}} \left\| \underset{\sim}{\alpha} - \underset{\sim}{\theta} \right\|^2$$

so that Condition (vii) always holds in some region about $\underset{\sim}{\alpha} = \underset{\sim}{\theta}$. That it hold throughout A requires that $G(\underset{\sim}{\alpha}) \, \underset{\approx}{m}(\underset{\sim}{\alpha})$ represent the gradient of a convex surface. We will consider this condition relative to an example in Section 4.

We now proceed to show that Assumptions 1 - 3 and 5 of Section 2 are satisfied.

First we remark that Assumption 1 follows directly from Conditions (iii) and Conditions (i) and (iv). Assumption 2 has been directly restated as Condition (ii). Assumption 3 follows immediately from Equation (3.25) and Condition (v).

Assumption 5, requires a little manipulation. Noting that $\underset{\approx}{m}(\underset{\sim}{\theta}) = \underset{\sim}{0}$, we have

$$(3.31) \quad \sum_{j=1}^{M} g_{ij}(\underset{\sim}{\alpha}) \, m_j(\underset{\sim}{\alpha}) = \sum_{j,k=1}^{M} \left[g_{ij}(\underset{\sim}{\alpha}) \frac{\partial}{\partial \alpha_k} m_j(\underset{\sim}{\alpha}) \right]_{\underset{\sim}{\alpha} = \underset{\sim}{\theta}} (\alpha_k - \theta_k)$$

$$+ \frac{1}{2} \sum_{j,k,\ell=1}^{M} \left[\frac{\partial^2}{\partial z_k \partial z_\ell} g_{ij}(\underset{\sim}{z}) \, m_j(\underset{\sim}{z}) \right] \Bigg|_{\substack{z = \underset{\sim}{\theta} + \beta(\underset{\sim}{\alpha} - \underset{\sim}{\theta}) \\ 0 < \beta < 1}} (\alpha_k - \theta_k)(\alpha_\ell - \theta_\ell) \, .$$

Upon applying Equations (3.25) and (3.30) and Condition (vi) we have

$$(3.32) \qquad \sum_{j=1}^{M} g_{ij}(\underset{\sim}{\alpha})\, m_j(\underset{\sim}{\alpha}) = -(\alpha_i - \theta_i) + \zeta_i$$

in which $\quad |\zeta_i| \le (K_1/2\sqrt{M})\, \|\underset{\sim}{\alpha} - \underset{\sim}{\theta}\|^2$

whence

$$(3.33) \qquad G(\underset{\sim}{\alpha})\, \underset{\sim}{m}(\underset{\sim}{\alpha}) = -(\underset{\sim}{\alpha} - \underset{\sim}{\theta}) + \zeta$$

with $\quad \|\underset{\sim}{\zeta}\| \le (K_1/2)\, \|\underset{\sim}{\alpha} - \underset{\sim}{\theta}\|^2$.

Lastly, using Equation (3.30)

$$(3.34) \qquad E\left\{\underset{\sim}{Y}'(\underset{\sim}{\alpha})\, G'(\underset{\sim}{\alpha})\, G(\underset{\sim}{\alpha})\, \underset{\sim}{Y}(\underset{\sim}{\alpha})\right\} = \sum_{i,j,k=1}^{M} E\left\{Y_i(\underset{\sim}{\alpha})\, Y_j(\underset{\sim}{\alpha})\right\} g_{ji}(\underset{\sim}{\alpha}) g_{jk}(\underset{\sim}{\alpha})$$

$$= \sum_{i,j,k=1}^{M} b_{ik}(\underset{\sim}{\alpha})\, g_{ij}(\underset{\sim}{\alpha})\, g_{jk}(\underset{\sim}{\alpha})$$

$$= \sum_{i,j,k=1}^{M} b_{ik}(\underset{\sim}{\theta})\, g_{jk}(\underset{\sim}{\theta})\, g_{ij}(\underset{\sim}{\theta})$$

$$+ \sum_{\ell=1}^{M} \frac{\partial}{\partial z_\ell}\left[\sum_{i,j,k=1}^{M} b_{ik}(\underset{\sim}{z})\, g_{ij}(\underset{\sim}{z})\, g_{jk}(\underset{\sim}{z})\right]\Bigg|\, (\alpha_\ell - \theta_\ell)$$

$$z = \underset{\sim}{\theta} + \beta(\underset{\sim}{\alpha} - \underset{\sim}{\theta})$$

$$0 < \beta < 1$$

$$= \sum_{k=1}^{M} g_{kk}(\underset{\sim}{\theta}) + \eta$$

in which $\quad |\eta| \le K_2\, \|\underset{\sim}{\alpha} - \underset{\sim}{\theta}\|$.

Thus our functions $G(\underset{\sim}{\alpha})$, $\underset{\sim\sim}{m}(\underset{\sim}{\alpha})$, and $B(\underset{\sim}{\alpha})$ satisfy Assumption 5.

Although the set of conditions set forth in this section makes certain stringent requirements on the existence of higher order derivatives, the first six conditions are essentially regularity conditions which can reasonably be assumed to hold for models representing those physical problems of greatest interest. Condition (vii), however, is another matter; it essentially deliniates which situations can be handled by the modified Robbins-Munro method. To show that the class of problems satisfying this condition is not vacuous, we consider a simple example in the next section.

4. Computational aspects; an example. In this section we wish to briefly comment on some of the computations required by the recursive estimation procedure and give consideration to Condition (vii) relative to a simple example.

The recursive procedure requires computation of the quantities

$$(3.28)\quad Y_{n,i}(\underset{\sim}{\alpha}) = \left(\frac{1}{2}\right)\int_0^T \int_0^T dt\, ds\, h^i(t,s,\underset{\sim}{\alpha}) \Big[W(t+\tau_n)\,W(s+\tau_n) - \phi_W(t,s,\underset{\sim}{\alpha})\Big]$$

$$i = 1, 2, \ldots, M .$$

Each of these quantities can be separated in the obvious way into two terms, only one of which depends on the received signal, $W(t)$. Using the fact that $h^i(t,s,\underset{\sim}{\alpha})$ is symetric in t and s (which follows from the symmetry of the covariance function), this term can be expressed in the form [6], [9]

$$\left(\frac{1}{N}\right)\int_0^T dt\, W(t) \int_0^t ds\; W(s)\, h^i(s,t,\underset{\sim}{\alpha}) \qquad i = 1, 2, \ldots, M$$

and thus can be computed in real time as the signal, $W(t)$, is received.

Note that our procedure depends upon being able to solve the integral Equation (3.6) for $h(r, t, \underset{\sim}{\alpha})$. In the case in which $S(t)$ is a sample of a stationary process with a rational spectral density, the solution to this equation is known [4], [12]. Moreover, our procedure is particularly useful in the case when the noise is much stronger than the signal (since then "good" estimation requires the processing of large amounts of data). In this case, it is usually true that

$$(4.1) \qquad \lambda_1(\underset{\sim}{\alpha}) \ll N$$

in which $\lambda_1(\underset{\sim}{\alpha})$ is the largest of the eigenvalues associated with $\phi_s(t, s, \underset{\sim}{\alpha})$. Thus in this situation, $h(t, s, \underset{\sim}{\alpha})$ is given to a good approximation by

$$(4.2) \qquad h(t, s, \underset{\sim}{\alpha}) \simeq \left(\frac{1}{N}\right) \phi_s(t, s, \underset{\sim}{\alpha}) \quad .$$

Lastly we consider an example, and investigate Condition (vii) relative to this example. Let $S(t)$ be a gaussian Markoff process of zero-mean and unknown variance and time constant. This process is observed in the presence of gaussian zero-mean "white" noise, so that

$$(4.3) \quad \phi_w(t, s, \underset{\sim}{\alpha}) = \delta(t - s) + \phi_s(t, s, \underset{\sim}{\alpha}) = \delta(t - s) + \alpha_1 \exp\left\{-\alpha_2 \left|t - s\right|\right\}$$

in which we have taken $N = 1$ for convenience. Our objective is to measure θ_1 and θ_2, the values of α_1 and α_2 actually pertaining to the process being observed. For this case the function $h(t, s, \underset{\sim}{\alpha})$ is known explicitly [7], [9]. However, in most cases of physical interest, the noise is much stronger than the signal; i.e.,

$$\frac{\alpha_1}{\alpha_2} << 1 \tag{4.4}$$

and the observation interval T is usually long compared to the correlation time of the process; i.e.,

$$\alpha_2 T >> 1 \quad . \tag{4.5}$$

The conditions expressed by inequalities (4.4) and (4.5) imply that the condition stated by inequality (4.1) holds, and hence that the approximation given by Equation (4.2) is good. For computational convenience, we assume that these two conditions are met and use the approximation of Equation (4.2) in investigating Condition (vii).

Using this approximation, we obtain (after much calculation) the approximate relation

$$G(\underset{\sim}{\alpha}) \underset{\sim}{m} (\underset{\sim}{\alpha}) = \left. \begin{array}{c} \dfrac{(\theta_1 - \alpha_1) \;\; \theta_2^2 + \left[\theta_2(\alpha^2 - \theta_2)\right] - \alpha_1(\alpha_2 - \theta_2)}{\left[\theta_2 + (\alpha_2 - \theta_2/2\right]^2} \\ \\ \\ \dfrac{(\theta_2 - \alpha_2) \;\; \theta_1 \;\; (\alpha_2)^2}{\left[\theta_2 + (\alpha_2 - \theta_2/2\right]^2} \end{array} \right] \tag{4.6}$$

so that

$$(4.7) \qquad -(\underset{\sim}{\alpha} - \underset{\sim}{\theta})\, G(\underset{\sim}{\alpha})\, \underset{\sim\sim}{m}\, (\underset{\sim}{\alpha}) = (\alpha_1 - \theta_1)^2\, \theta_2 \alpha_2 \left[\theta_2 + (\alpha_2 - \theta_2)/2\right]^{-2}$$

$$+ (\alpha_2 - \theta_2)^2 \left[\theta_1 (\alpha_2)^2/\alpha_1 + \alpha_1(\alpha_1 - \theta_1)/4\right]$$

$$\cdot \left[\theta_2 + (\alpha_2 - \theta_2)/2\right]^{-2} .$$

Thus (noting that by assumption $\alpha_2/\alpha_1 >> 1$) as long as the search is confined to strictly positive values

$$0 < \epsilon_1 \leq \alpha_1 \; ; \qquad\qquad 0 < \epsilon_2 \leq \alpha_2$$

Condition (vii) will be satisfied.