p-EFFICIENT ESTIMATORS

by

P. PAPANTONI-KAZAKOS

AUGUST 1975

TECHNICAL REPORT #7512
p-EFFICIENT ESTIMATORS

by

P. Papantoni-Kazakos
Rice University
Houston, Texas

ABSTRACT

This paper is concerned with estimates of an unknown vector parameter \( S \) based on observations \( X \). A generalized error autocorrelation matrix with components the error moments of order \( 2p \), is defined. A lower bound for this matrix is found and the p-efficient \( P(X|S) \) statistics realizing it are determined. The cases of i) \( X \) and \( S \) real, ii) \( S \) scalar real, and iii) \( X \) real and \( S \) complex are examined.

1. INTRODUCTION AND SUMMARY OF RESULTS

Let us assume that the estimation of a \((k+1)\)-dimensional parameter \( S \) is the subject of interest. Available to the investigator is the \( m \)-dimensional observation vector \( X \) and a measure of evaluation of the different estimators \( \hat{S}(X) \) is sought.

One such measure that has been widely used until now is the error autocorrelation matrix \( R(\hat{S},S) \) which is given by the following expression

\[

(1)

A lower bound for the above matrix and the efficient estimators that realize it have been established (Cramér-Rao bound)[1]. The matrix in (1) has as components the second-order crossmoments of the components of \( \hat{S}(X)-S \) and its trace is equal to the mean square error \( E([\hat{S}(X)-S]^T[\hat{S}(X)-S]^*/S) \). In many real cases, though, one is also interested in error moments of order

\[ (1) \text{ This research was sponsored in part by the National Science Foundation under Grant Number: ENG 75-10516.} \]
higher than two. Such is the case of $S(X)-s$ having non-Gaussian statistics. Then, all the even error moments are theoretically needed (practically a few only) for the evaluation of the estimate $\hat{S}(X)$. The odd error moments are not so interesting, since for symmetric distributions at least, they are all equal to zero.

The subject of this paper will be the obtaining of a lower bound for a generalized $2p$-order error autocorrelation matrix which will be defined below. The estimators as well as the P(X/S) statistics that realize this lower bound will be studied. Such estimators will be called $p$-efficient. In section 2, the case of both $X$ and $S$ being real is examined and the $p$-efficient estimators are defined. In section 3, the case of $S$ being scalar is studied separately. In section 4, the case of $X$ real is $S$ complex is presented.

2. X AND S BOTH REAL

In this case the arguments $X$, $S$ and $\hat{S}(X)$ involved are all real. The differential operator $\nabla$ will be first defined as a column vector of dimensionality $k+1$ and such that the $i$th component of $\nabla(\cdot)$ is equal to $\partial(\cdot)/\partial s_i$ where

$$S = \{s_i, i = 0, \ldots, k\} \tag{2}$$

The following theorem can then be stated.

Theorem 1

For P(X/S) statistics and $\hat{S}(X)$ estimates, such that the following two equalities are true

$$\nabla\int P(X/S) dX = \int \nabla P(X/S) dX \tag{3}$$

$$\nabla\int S^T(X)P(X/S) dX = \int \nabla P(X/S) S^T(X) dX \tag{4}$$
and for arbitrary (k+1)-dimensional column constant vector \( A \) and \((k+1)\times(k+1)\) square matrix \( M(S) \), which is in general a function of the true value of the parameter \( S \), the following inequality is satisfied

\[
\frac{1}{E^D(A^T M(S) [S(X) - S])^{2p/S}} \geq \frac{1}{\int dX |A^T \nabla \log P(X/S)|^{2p/(2p-1)} P(X/S)^{(2p-1)/p}}
\]

(5)

\( p \) is an arbitrary natural number and (5) is satisfied with equality if and only if

\[
A^T \nabla \log P(X/S) = C[A^T M(S) [S(X) - S]]^{2p-1}
\]

(6)

where \( C \): arbitrary constant

\( A^T \): the transpose of the vector \( A \)

**Proof**

It is true that for any \( P(X/S) \) distribution, integration with respect to \( X \) over the whole range, gives one. That is

\[
\int P(X/S) dX = 1
\]

(7)

Using the operator \( \nabla \) , as defined in this section, we obviously have from (7)

\[
\nabla \int P(X/S) dX = 0
\]

(8)

Applying (3) to (8), we conclude that if (3) is satisfied, so is the following equality

\[
\int dX \cdot \nabla P(X/S) = 0
\]

(9)

Let us now use the arbitrary vector \( A \) and the matrix \( M(S) \) that were introduced in the statement of this theorem. Then, because of the obvious equality
\[ \forall \mathbf{E}^T \{ \mathbf{S}(X)/S \} = \int \mathbf{S}(X) \nabla^T \mathbf{P}(X/S) dX \quad (10) \]

we obtain

\[ \mathbf{A}^T \mathbf{M}(S) \forall \mathbf{E}^T \{ \mathbf{S}(X)/S \} \mathbf{A} = \int \mathbf{A}^T \mathbf{M}(S) \mathbf{S}(X) \nabla^T \mathbf{P}(X/S) dX \quad (11) \]

If (9) is applied to (11), one can directly write

\[ \mathbf{A}^T \mathbf{M}(S) \forall \mathbf{E}^T \{ \mathbf{S}(X)/S \} \mathbf{A} = \int \mathbf{A}^T \mathbf{M}(S) \mathbf{S}(X) \mathbf{A}^T \nabla \mathbf{P}(X/S) dX - \mathbf{A}^T \mathbf{M}(S) \mathbf{S} \mathbf{A}^T \int \nabla \mathbf{P}(X/S) dX = \]

\[ = \int \mathbf{A}^T \mathbf{M}(S) [\mathbf{S}(X)-\mathbf{S}] \mathbf{A}^T \nabla \mathbf{P}(X/S) dX \quad (12) \]

We can now express the vector \( \nabla \mathbf{P}(X/S) \) in the following form

\[ \nabla \mathbf{P}(X/S) = \mathbf{P}(X/S) \cdot \nabla \log \mathbf{P}(X/S) = [\mathbf{P}(X/S)]^{1/2p} \cdot [\mathbf{P}(X/S)]^{1/q} \cdot \nabla \log \mathbf{P}(X/S) \quad (13) \]

In (13) \( p \) and \( q \) are real, positive and such that

\[ \frac{1}{2p} + \frac{1}{q} = 1 \quad (14) \]

For the interests of the present paper, \( p \) is also chosen to be a natural number. Substitution of (13) in (12) results in the following expression

\[ \mathbf{A}^T \mathbf{M}(S) \forall \mathbf{E}^T \{ \mathbf{S}(X)/S \} \mathbf{A} = \int dX \mathbf{A}^T \mathbf{M}(S) [\mathbf{S}(X)-\mathbf{S}] \mathbf{P}^{2p}(X/S) [\mathbf{A}^T \nabla \log \mathbf{P}(X/S) \mathbf{P}^q(X/S)] \quad (15) \]

Let us now examine the following different version of the right-hand side of (15).

\[ \int dX \frac{1}{[\mathbf{A}^T \mathbf{M}(S)[\mathbf{S}(X)-\mathbf{S}] \mathbf{P}^{2p}(X/S)] [\mathbf{A}^T \nabla \log \mathbf{P}(X/S) \mathbf{P}^q(X/S)]} \quad (16) \]

The difference of expression (16) from (15) is the absolute values in the integrand of the first one.

Applying the Hölder inequality on (16), we obtain
\[
\frac{1}{\int \left| A^T M(S)[S(X)-S] P^2 P(X/S) \right| \left| A^T \nabla \log P(X/S) P^q P(X/S) \right|} \leq \frac{1}{\left( \int \left| A^T M(S)[S(X)-S] \right|^2 P^2 P(X/S) \right)^{2p} \left( \int \left| A^T \nabla \log P(X/S) \right|^q P(X/S) \right)^{q}}
\]

It is well known that (17) holds with equality if and only if
\[
\left| A^T M(S)[S(X)-S] \right|^{2p} = C_0 \left| A^T \nabla \log P(X/S) \right|^q
\]
for almost every \( X \) where \( C_0 \) is an arbitrary positive constant. Since the expressions in (18) are all scalar and real, one directly gets from it:
\[
A^T \nabla \log P(X/S) = C \left[ A^T M(S)[S(X)-S] \right]^q
\]
(19)

Because of (14), (19) changes to:
\[
A^T \nabla \log P(X/S) = C \left[ A^T M(S)[S(X)-S] \right]^{2p-1}
\]
(20)
where \( C \) is now an arbitrary (either positive or negative) real constant.

At this point, let us observe that
\[
\frac{1}{\int \left| A^T M(S)[S(X)-S] P^2 P(X/S) \right| \left| A^T \nabla \log P(X/S) P^q P(X/S) \right|} \leq \frac{1}{\int \left| A^T M(S)[S(X)-S] \right|^2 P^2 P(X/S) \left| A^T \nabla \log P(X/S) \right|^q P(X/S) \left| \right|^q}
\]

Applying (21) to (15) and (17) we get
\[
\left| A^T M(S) \nabla E(T)(S(X/S)A) \right| \leq \left[ \int \left| A^T M(S)[S(X)-S] \right|^2 P^2 P(X/S) \right]^{2p} \left[ \int \left| A^T \nabla \log P(X/S) \right|^q P(X/S) \right]^{q} = \left[ \int \left| A^T M(S)[S(X)-S] \right|^2 P^2 P(X/S) \right]^{2p} \left[ \int \left| A^T \nabla \log P(X/S) \right|^{2p-1} P(X/S) \right]^{2p-1}
\]
(22)

In the development of (22), the expression (14) was used. It was shown that (17) holds with equality if and only if (20) is true. It can easily be shown
now that if (20) holds (21) is true with equality. Therefore, (22) holds with equality if and only if (20 is true. If the second power of both parts of the inequality in (22) is taken and the two parts of the resulting inequality are then divided by

$$
\frac{1}{\int dX \left[ A^T M(S) \left[ S(X) - S \right]^2 \right]^{2p} P(X/S)} \frac{2p}{\int dX \left[ A^T \log P(X/S) \right]^{2p-1} P(X/S)}
$$

one obtains the inequality (15) of the statement of the theorem with the if and only if condition expressed by (6). The following relationship was used in the process

$$
\int dX \left[ A^T M(S) \left[ S(X) - S \right]^2 \right]^{2p} P(X/S) = E \left[ \left[ A^T M(S) \left[ S(X) - S \right]^2 \right]^{2p} / S \right]
$$

(23)

In relationship (5), it is not clear what the expectation \( E[ A^T M(S) \left[ S(X) - S \right]^{2p} / S \] represents for the arbitrary \( p \). If \( p = 1 \), the above expectation is clearly a transformation of the error autocorrelation function of \( S(X) - S \), (5) is then the Cramér-Rao information inequality and (6) the necessary and sufficient condition for (5) to be equality.

We will use a few definitions and transformations now to express (5) and (6) in a more meaningful way. For that, the definitions will first be given and then the new changed versions of (5) and (6) will be stated by a theorem.

Let us define

$$
B(S) = M^T(S) \cdot A = \{ b_i ; i = 0, \ldots, k \}
$$

(24)

$$
B_p = \left\{ \frac{p!}{1} i_0 \ldots i_k \right\} a^{i_0} \ldots b_k \cdot \sum_{m=0}^{m=p} i_m
$$

(25)
\[ S(X) - S = \{ s_{d_i} ; i = 0, \ldots, k \} \]  \hspace{1cm} (26)

\[ S_p(X) = \{ s_{d_0} \ldots s_{d_k} ; \Sigma i_m = p \} \]  \hspace{1cm} (27)

\[ vE^{T}(S(X)/S)A = \{ d_i ; i = 0, \ldots, k \} \]  \hspace{1cm} (28)

\[ D_p(S) = \{ d_{d_0} \ldots d_{d_k} ; \Sigma i_m = p \} \]  \hspace{1cm} (29)

In definitions (24-29) the expressions in the brackets indicate the general component of the corresponding defined vector. Based on the above definitions, the following theorem can be stated.

**Theorem 2**

For the arbitrary natural number \( p \), \((k+1)\)-dimensional constant vector \( A \) and \((k+1)\)-dimensional vector \( B(S) \), as well as \( P(X/S) \) statistics satisfying (3) and (4), the following inequality is satisfied:

\[
\frac{B^T(S)E(S(X)S^T(X)/S)B(S)}{B^T(S)D(S)D^T(S)B(S)} \geq \frac{1}{\frac{2p}{\int [dX|x^T\log P(X/S)|^{2p-1}P(X/S)]^{2p-1}}} 
\]

\[ T \]

In (30), the vectors defined in (24-29) were used. Furthermore, (30) is satisfied with equality if and only if

\[
\frac{A^T}{\|A\|} v \log P(X/S) = CB^T_{2p-1}(S)S_{2p-1}(X) 
\]

(31)

where \( C \) is an arbitrary real constant.
Proof

Basically, the relationships (5) and (6) in the statement of Theorem 1 will be expressed in a way involving the definitions (24-29).

Let us start from the $2p$ power of the inner product $A^T M(S)[S(X)-S]$. Due to (24) and (26), the following equation is obtained

$$[A^T M(S)[S(X)-S]]^{2p} = \left[ \sum_{i=0}^{k} b_i s_{d_i} \right]^{2p}$$

(32)

Expanding the right hand part of (32), we get

$$[A^T M(S)[S(X)-S]]^{2p} = \left[ \sum_{i=0}^{k} b_i^{o} \cdots b_i^{k} s_{d_i}^{o} \cdots s_{d_i}^{k} i_1! \cdots i_k! \right]^{2p}$$

(33)

In the summation in (33), the number of terms is equal to $\binom{k+p}{p}$, and the sum is the inner product of the vectors $B_p$ and $S(X)$ which are defined by (25) and (27) correspondingly. Therefore, one can express (33) as follows

$$[A^T M(S)[S(X)-S]]^{2p} = [B_p^T S(X)]^2 = B_p^T S(X) S_p^T (X) B_p (S)$$

(34)

Taking the mean of both parts of (34), one obtains

$$E([A^T M(S)[S(X)-S]]^{2p}/S) = B_p^T S(X) S_p^T (X)/S B_p (S)$$

(35)

In a similar way we obtain

$$[A^T M(S) \vee E [S(X)/S] A]^{2p} = \left[ \sum_{i=0}^{k} b_i d_i \right]^{2p} = \left[ \sum_{i=0}^{k} b_i^{o} \cdots b_i^{k} d_{o} \cdots d_{k} i_1! \cdots i_k! \right]^{2p}$$

(36)

where $i_j \geq 0$, $\forall j$
\[ = [B^T_p(S)D_p(S)B^T_p(S)]^2 = B^T_p(S)D_p(S)D^T_p(S)B_p(S) \]  

(36)

In the development of (36), the definitions (24), (25), (28), and (29) were used.

Finally, using definitions (24), (25), (26) and (27) we also obtain

\[ \left[ A^T M(S)[S(X)-S] \right]^{2p-1} = \left[ \sum_{i=0}^{k} b_i d_i \right]^{2p-1} = \sum_{i=0}^{k} b_i^{i+1} s_i^{i} \cdot \sum_{j=0}^{k} i! \cdots i! = \sum_{i,j=0}^{k} i! \cdots i! \]

\[ = B^T_{2p-1}(S) \cdot S_{2p-1}(X) \]  

(37)

In inequality (5) all quantities involved are positive, therefore no change to it or the condition (6) will occur if the \( p \) power of both parts is taken. If this is done, expressions (35), (36) and (37) are used and both parts of (6) are divided by the \( ||A|| \) of \( A \), the inequality (30) and the equation (31) of the statement of theorem 2 are obtained. The proof of the theorem is now complete.

At this point, we consider important to make some observations:

1. The expectation \( E[S(X)S^T(X)/S] \) is a \( (p+k) \times (p+k) \) Hermitian matrix. It can be considered as a generalized error autocorrelation matrix, its components being the crossmoments

\[ \sum_{i,j=0}^{k} s_i^{i+j} \cdot s_j^{i+k} \]

of the \( s_i \) components of the vector \( S(X)-S \). These moments are of order \( 2p(\sum i \cdot j = p) \).
The expectation \( E[S(X)S^T(X)/S] \) will then be denoted \( R_p(S) \) and it will be called \textit{p-order conditional error autocorrelation matrix} of the estimate \( \hat{S}(X) \).

Due to the above, we might point out here that inequality (30) provides a lower bound on the \( p \)-order conditional error autocorrelation matrix while (31) provides the general family of statistics that satisfy it.

2. If \( p=1 \) and \( M(S) \) is either equal to the identity matrix or to the inverse of the first-order conditional error autocorrelation matrix \( R(S) \), inequality (30) becomes the widely known Cramér-Rao one and (31) provides the general family of \( P(X/S) \) statistics satisfying it with equality.

The estimates that satisfy the Cramér-Rao bound (give inequality in (30)) have been called efficient. As an extension to that, the following definition is given:

\textbf{Definition:} 

The estimates that satisfy the relationship (30) at the statement of theorem 2 with equality will be called \textit{\( p \)-efficient}.

3. Let us consider here the family of unbiased estimates. The vector with components \( d_i \) in (28) is then equal to the vector \( A \). If, in addition, the matrix \( M(S) \) is chosen to be independent of the estimate, the only expression in (30) including it will be, of course, the quadratic form

\[ B^T(S)R_p(S)B_p(S) \]

(38)

which then will be bounded from below by
\[
\frac{\mathbf{B}^T \mathbf{R} \mathbf{B}}{\mathbf{p}^{2p}} \frac{1}{\left[ \int \! \! \! \int \! \! \! \int \! \! \! \int \! \! \! \int A^T \partial \log \mathbf{P}(\mathbf{X}/\mathbf{S}) |^{2p-1} \mathbf{p}(\mathbf{X}/\mathbf{S}) \right]^{2p-1}}
\]

(39)

If now \( \mathbf{M}(\mathbf{S}) \) is chosen equal to the identity matrix, both vectors \( \mathbf{B}_p \) and \( \mathbf{D}_p \) are equal and they involve only \( p \)-order powers of the \( \mathbf{A} \) components. In this case the lower bound of (38) for the family of unbiased estimates becomes equal to

\[
\frac{\| \mathbf{B}_p \|^{2p}}{\left[ \int \! \! \! \int \! \! \! \int \! \! \! \int A^T \partial \log \mathbf{P}(\mathbf{X}/\mathbf{S}) |^{2p-1} \mathbf{p}(\mathbf{X}/\mathbf{S}) \right]^{2p-1}}
\]

(40)

4. Consider here the necessary and sufficient condition (31) for the relationship (30) to be equality. We can write

\[
\mathbf{B}^T_{2p-1} \mathbf{S}_{2p-1}(\mathbf{X}) = \left[ \sum_{i=0}^{k} \mathbf{b}_i \mathbf{s}_{d_i} \right]^{2p-1} = \left[ \sum_{i=0}^{k} \mathbf{b}_i \mathbf{s}_{d_i} \right] \left[ \sum_{i=0}^{k} \mathbf{b}_i \mathbf{s}_{d_i} \right]^{2p-2} = \\
= \mathbf{B}^T(\mathbf{S}(\mathbf{X}) - \mathbf{S}) \mathbf{B}^T_{2p-2} \mathbf{S}_{2p-2}(\mathbf{X}) = \mathbf{A}^T \mathbf{M}(\mathbf{S})[\mathbf{S}(\mathbf{X}) - \mathbf{S}] \mathbf{B}^T_{2p-2} \mathbf{S}_{2p-2}(\mathbf{X})
\]

(41)

Application of (41) to (31) gives

\[
\mathbf{A}^T \left[ \partial \log \mathbf{P}(\mathbf{X}/\mathbf{S}) - C_1 \mathbf{M}(\mathbf{S})[\mathbf{S}(\mathbf{X}) - \mathbf{S}] \mathbf{B}^T_{2p-2} \mathbf{S}_{2p-2}(\mathbf{X}) \right] = 0
\]

(42)

where

- \( C_1 \): arbitrary constant
- \( \mathbf{M}(\mathbf{S}) \): arbitrary matrix
- \( \mathbf{A} \): arbitrary vector
- \( \mathbf{B}_{2p-2}(\mathbf{S}) \): arbitrary vector

If we want (42) satisfied for any vector \( \mathbf{A} \), we should have

\[
\partial \log \mathbf{P}(\mathbf{X}/\mathbf{S}) = C_1 \mathbf{M}(\mathbf{S})[\mathbf{S}(\mathbf{X}) - \mathbf{S}] \mathbf{B}^T_{2p-2} \mathbf{S}_{2p-2}(\mathbf{X})
\]

(43)
Now, since \( B_{2p-2}(S)S_{2p-2}(X) \) is an inner product, we may change its order at (43) and get

\[
\nabla \log P(X/S) = C_1 B_{2p-2}^T(S)S_{2p-2}(X) \cdot M(S)[\hat{S}(X)-S] \tag{44}
\]

If we choose \( M(S) = I \), then \( B_{2p-2}(S) \) is an arbitrary constant vector and including the constant \( C_1 \) in it we get

\[
\nabla \log P(X/S) = B_{2p-2}^T \cdot S_{2p-2}(X)[\hat{S}(X)-S] \tag{45}
\]

Integrating (45) with respect to \( S \) one gets

\[
\log P(X/S) = B_{2p-2}^T \nabla^{-1} S_{2p-2}(X)[\hat{S}(X)-S]+K(X) \tag{46}
\]

where \( \nabla^{-1} \) indicates integration with respect to \( S \), \( B_{2p-2} \) is an arbitrary constant vector of dimensionality \( ^{2p+k-2}_{2p-2} \) and \( K(X) \) is an arbitrary function of \( X \) only. Directly from (46) we obtain

\[
P(X/S) = F(X) \cdot \exp[B_{2p-2}^T \nabla^{-1} S_{2p-2}(X)[\hat{S}(X)-S]] \tag{47}
\]

We will now summarize the part 4 of our observations by stating the following lemma.

**Lemma 1**

The most general family of \( P(X/S) \) statistics that satisfy (30) with equality, when \( B(S) = A \) (\( B(S) \) defined by (24)) is the following

\[
P(X/S) = F(X) \cdot \exp[B_{2p-2}^T \nabla^{-1} S_{2p-2}(X)[\hat{S}(X)-S]] \tag{48}
\]

The statistics described by (48) are exponential in a wide sense.

\( \hat{S}(X) \) is an \( X \) function that represents the p-efficient estimate in this case, \( B_{2p-2} \) is an arbitrary constant column vector and \( F(X) \) the function
that normalizes $P(X/S)$.

It is interesting to examine the one-dimensional ($S$ scalar) separately.

3. X AND S REAL, S SCALAR

This is the case that $k = 0$. Then, the vectors in (24-28) are reduced to scalars so that one can write

\[ \hat{S}(X) - S = s_d \]  
(49)

\[ B(S) = b_o \]  
(50)

\[ \forall E^T[\hat{S}(X)/S]A = d_o = \frac{\theta}{\theta_S} E[\hat{S}(X)/s] \]  
(51)

\[ B_p = b_p^o \]  
(52)

\[ S_p(X) = s_p^d \]  
(53)

\[ D_p(S) = d_p^o \]  
(54)

\[ \mu_{2p}(s) = E[(\hat{S}(X) - \cdot)^{2p}/s] \]  
(55)

Application of definitions (49-55) in (30) and (31) leads to the following lemma.

Lemma 2

If $p$ is an arbitrary natural number and $\hat{S}(X)$ an arbitrary estimate of the scalar parameter $s$ based on the $m$-dimensional observation vector $X$, then the $2p$-order conditional error moment $\mu_{2p}(s)$ of the estimate $\hat{S}(X)$ is bounded from below by a quantity expressed by the following inequality.
\[ \mu_{2p}(s) \geq \frac{\left[ \frac{\theta}{\delta s} E[\hat{s}(X)/s] \right]^{2p}}{\left[ \int dX \left[ \frac{\theta}{\delta s} \log P(X/s) \right]^{2p-1} P(X/s) \right]^{2p-1}} \]  

(56)

The most general family of statistics satisfying (56) with equality is expressed by the following relationship

\[ P(X/s) = K(X) \exp \left[ \int ds \cdot C(s)[\hat{s}(X)-s]^{2p-1} \right] \]  

(57)

where \( C(s) \) is an arbitrary function of \( s \), \( K(X) \) the normalizing function of \( X \) and \( \hat{s}(X) \) the estimate that makes the realization of the lower bound in (56) possible.

For the class of unbiased estimates, the moment \( \mu_{2p}(s) \) is bounded from below by

\[ \frac{1}{\left[ \int dX \left[ \frac{\theta}{\delta s} \log P(X/s) \right]^{2p-1} P(X/s) \right]^{2p-1}} \]  

(58)

We may observe here that from the expression (57) of the general distribution families that realize the bound in (56), a subclass can be extracted. This is the subclass characterized by a function \( C(s) \) which is a constant with respect to \( s \). The \( P(X/s) \) statistics of this class are obviously given by the following expression

\[ P(X/s) = K(X) \cdot \exp [C[\hat{s}(X)-s]^{2p}] \]  

(59)

(59) describes an exponential distribution family which for \( p = 1 \) coincides with the class that satisfies the Cramér-Rao bound.
4. SEMI-COMPLEX CASE

In this section the case of the observations \( X \) being real but the parameter \( S \) being complex will be considered. The reason this problem is examined separately and is not considered a subcase of section 2 is that our effort is directed toward keeping the dimensionality of the problem equal to the one of the parameter \( S \).

Our analysis here will be useful in the case that emphasis is only on the development of a lower bound on a p-order conditional error autocorrelation matrix and not on the statistics that realize this bound.

The differential operator \( \nabla \) will now be defined as a \((k+1)\)-dimensional column vector such that the \( i \)th component of \( \nabla(\cdot) \) is equal to

\[
\frac{\partial}{\partial s_i} + j \frac{\partial}{\partial s_{i2}}
\]

\( s_{i1} \) and \( s_{i2} \) are correspondingly the real and imaginary part of the \( i \)th component of the parameter vector \( S \).

Let us also define, similarly to definitions (24-29), the following vectors

\[ A : \text{arbitrary} \ (k+1)\text{-dimensional complex column vector} \]

\[ B(S) \text{ and } B(S) \text{ as defined by (24) and (25).} \]

\[
[\hat{S}(X)-S]^* = (s_{di}, i = 0, \ldots, k) \\
S_p^*(X) = \{s_{di}^{(0)} \ldots s_{di}^{(k)} ; \sum_{m=0}^{k} s_{di}^{(m)} = p\}
\]

\[
\nabla E^T[S(X)/S]A^* = (d_i ; i = 0, \ldots, k)
\]
\[
D_p^*(S) = \{d_0^* \ldots d_k^* ; \sum_{m=0}^{k} i_m = p \}\tag{63}
\]

\[
E(X/S) = \nabla \log P(X/S) = \{e_i ; i = 0, \ldots, k\}
\]

\[
E_p^*(X/S) = \{e_0^* \ldots e_k^* ; \sum_{m=0}^{k} i_m = p\}
\]

where \((\quad)^*\) means conjugate. Then, the following lemma can be stated.

**Lemma 3**

If \(p\) is an arbitrary natural number and the \(P(X/S)\) statistics as well as the estimate \(S(X)\) are such that

\[
\nabla \int dX P(X/S) = \int dX \nabla P(X/S) = \int dX \nabla P(X/S) S^*(X) = \int dX \nabla P(X/S) S^*(X)\tag{67}
\]

then for some arbitrary \(A\) and \(B(S)\), the following inequality is true:

\[
\frac{B^T(p) E[S^*(X) \cdot S^*(X)] B^*(p)}{B^T(p) D^*(p) D^*(p) B^*(p)} \geq \frac{1}{\int \int dX \|A^T \nabla \log P(X/S)\|^2 p-1 P(X/S)^{2p-1}}\tag{68}
\]

where the notations appearing in (68) have been defined in (60-65) and \(|\quad|\) means norm.

**Proof**

Following a procedure similar to the one of theorems 1 and 2, we have

\[
A^T M(S) \nabla^T S^*(X) A^* = \int dX A^T M(S) S^*(X) \nabla^T P(X/S) A^*
\]

Using (67) in the above equality we obtain
\[ A^T M(S) \forall E (\mathbf{s}^*(X)/S) A^* = \int dX A^T M(S)[\mathbf{s}(X)-S]^* A^T \nabla P(X/S) = \frac{1}{\int dX [A^T M(S)[\mathbf{s}(X)-S]^* P^{2p}(X/S)] [A^T \nabla \log P(X/S) P^q(X/S)]} \]

where \( \frac{1}{2p} + \frac{1}{q} = 1 \)

and \( p \) and \( q \) are real positive numbers.

Applying Hölder inequality on (69) we obtain

\[ \frac{1}{\int dX [A^T M(S)[\mathbf{s}(X)-S]^* ||P^{2p}(X/S) - ||A^T \nabla \log P(X/S)||^q P(X/S) \leq \int dX [A^T M(S)[\mathbf{s}(X)-S]^* ||^2 P^{2p}(X/S)]^2 \cdot [\int dX ||A^T \nabla \log P(X/S)||^q P(X/S)]^q \]

with equality if and only if

\[ ||A^T M(S)[\mathbf{s}(X)-S]^* ||^2 P^{2p} = C ||A^T \nabla \log P(X/S)||^q \]

(71)

where \( C \) is an arbitrary real positive constant.

At the same time, while it is true that

\[ ||A^T M(S) \forall E (\mathbf{s}^*(X)/S) A^* || \leq \int dX [A^T M(S)[\mathbf{s}(X)-S]^* ||P^{2p}(X/S) - ||A^T \nabla \log P(X/S)||^q P(X/S) \]

(72)

it is not clear whether or not (71) satisfies (72) with equality or not.

Therefore, the best that one can do here is express the following relationship

\[ ||A^T M(S) \forall E (\mathbf{s}^*(X)/S) A^* || \leq \left[ \int dX [A^T M(S)[\mathbf{s}(X)-S]^* ||^2 P^{2p}(X/S)] \right]^{2p} \cdot \left[ \int dX ||A^T \nabla \log P(X/S)||^q P(X/S) \right]^q \]

(73)

Taking the \( 2p \)-power of both parts of (73) and substituting \( q \) by \( \frac{2p}{2p-1} \),

we get
\[
\frac{E\left(\|A^T M(S)[S(X)-S]^*_p\|^2p\right)}{\|A^T M(S)\vee E\left(S^*_p(X)/S\right)A^*_p\|^2p} \geq \frac{1}{\left[\int dX \|A^*T \vee \log P(X/S)\|^2p-1 P(X/S)\right]^{2p-1}}
\]

(74)

If definitions (60-63) are used in (74), one obtains inequality (68) of the lemma. Conditions for which (68) is satisfied with equality cannot be derived. The proof of the theorem is now complete.

Here we can observe again that (68) provides a bound on the p-order conditional error autocorrelation matrix

\[
R_p(S) = E(S^*_p(X)S^T_p(X))
\]

Furthermore, if \( M(S) \) is chosen to be the identity matrix and we restrict ourselves to the class of unbiased estimates, we have a lower bound for \( R_p(S) \) expressed by the following relationship

\[
B^T R_p(S)B^*_p \geq \frac{\|B_p\|^4}{\left[\int dX [B^T E(X,S)E^T(X,S)B^*_p]^{2p-1} P(X/S)\right]^{2p-1}}
\]

(75)

In (75) definitions (64) and (65) were also used.
REFERENCES

