ML ESTIMATION OF SIGNAL POWER IN THE PRESENCE OF UNKNOWN NOISE FIELD – SIMPLE APPROXIMATE ESTIMATOR AND EXPLICIT CRAMER–RAO BOUND

Alex B. Gershman *Alexander L. Matveyev* *Johann F. Böhme*

Institute of Applied Physics, Nizhny Novgorod, Russia

Ruhr University, Bochum, FRG

ABSTRACT

A simple approximate maximum likelihood (AML) estimator is derived for estimating a power of a single signal with rank-one spatial covariance matrix known *a priori* except for a scaling. The noises are assumed to have different and unknown powers in each array sensor.

The variance of the introduced AML estimator is compared with the exact Cramer-Rao bound (CRB) of this estimation problem analytically and by computer simulations. It is shown analytically that the AML estimator achieves CRB in the majority of practically important cases.

Computer simulations have been performed showing that the estimation errors of the AML estimator are very close to CRB for a wide SNR range.

1. INTRODUCTION

ML estimation of signal parameters from sensor array noisy data has received considerable attention [1], [2]. Usually, the total ML solutions are computationally expensive. Generally, when *a priori* information about the covariance matrix structure or about covariance components is taken into account, the ML estimator performance is greatly improved as well as yielding simpler implementations of the ML estimation.

In [3], [4] the problem of signal and noise power estimation has been considered for different cases of structured covariance. In [3] the simple ML estimator of signal and noise powers has been derived for the special case when the signal and noise covariance matrices have low rank and are known *a priori* except for a scaling. In [4] the same problem has been considered for the case of full rank noise covariance matrix and arbitrary rank signal covariance matrix.

In many practical situations in radar and sonar sensor noise powers may be different and unknown. In this paper we consider the problem similar to [3], [4] assuming that the powers of noise are arbitrary and unknown in each array sensor and that the noise is spatially uncorrelated. The spatial covariance matrix of received signal from single source is assumed to be a rank-one matrix and to be known *a priori* except for a scaling, i.e., except for a signal power. However, the results can be extended to the case of well separated multiple sources. We derive the simple approximate ML estimator (called below as AML) of signal power in the presence of unknown sensor noises assuming that the signal-to-noise ratio (SNR) is low and that the number of data snapshots is large. The variance of the AML estimator is compared analytically with the exact explicit CRB of this problem. Such a comparison allows us to prove that the derived estimator approaches asymptotic efficiency in several practically important cases.

Computer simulations show that the root-mean-square estimation errors of our estimator are very close to the CRB for a wide SNR range.

This work was supported by INTAS under SASPARC project and by Alexander von Humboldt Foundation under Research Fellowship.

2. AML ESTIMATOR

Assume that single signal impinges on the array of n sensors. Hence, a $n \times 1$ complex vector of array outputs can be expressed as [1], [2]:

$$\boldsymbol{x}(i) = \boldsymbol{a}s(i) + \boldsymbol{n}(i), \qquad i = 1, 2, \ldots, N \qquad (1)$$

where $\boldsymbol{a}$ is the $n \times 1$ signal wavefront vector (so-called steering vector), $s(i)$ is the random signal waveform, N is the total number of data snapshots. The spatial covariance matrix of the received signals (1) is defined as $\boldsymbol{M} = E\left[\boldsymbol{x}(i)\boldsymbol{x}^H(i)\right]$, where H denotes the Hermitian transpose and $E[\cdot]$ denotes the expectation.

Let us now make the following assumptions concerning the data vector model:

1). The data vector $\boldsymbol{x}(i)$ is the stationary, Gaussian complex random vector with the following properties:

$$E\left[\boldsymbol{x}(i)\right] = 0, \quad E\left[\boldsymbol{x}(i)\boldsymbol{x}^H(k)\right] = \delta_{ik}\boldsymbol{M}$$

where δ_{ik} denotes Kronecker delta.

0-7803-2431-5/95 $4.00 © 1995 IEEE

2). The steering vector $\boldsymbol{a}$ is known *a priori*.

3). The elements of the noise vector $\boldsymbol{n}(i)$ are statistically independent of each other and also statistically independent of $s(i)$.

4). The signal waveforms, $s(i)$ are stationary, statistically independent, and zero-mean complex Gaussian random quantities with unknown power p_S.

5). The noise powers σ_k^2, $k = 1, 2, \ldots, n$ are unknown.

Thus, the spatial covariance matrix has the form

$$\boldsymbol{M} = \boldsymbol{M}_N + p_S \boldsymbol{a}\boldsymbol{a}^H$$

where p_S is a signal power, $\boldsymbol{M}_N$ is a diagonal noise covariance matrix

$$\boldsymbol{M}_N = \operatorname{diag}\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2\}$$

In other words, according to these assumptions, the covariance matrix can be represented as a sum of two matrices. The first one (the noise covariance matrix) is known *a priori* to be a diagonal matrix but its diagonal elements (i.e., noise powers) are arbitrary and unknown. The second one is known *a priori* except for a scaling (i.e., except for a signal power p_S) rank-one signal covariance matrix. The problem is to derive the ML estimate $\hat{p}_S$ of signal power p_S.

Using *a priori* knowledge of steering vector $\boldsymbol{a}$, let us transform the array data (1) as follows:

$$\boldsymbol{y}(i) = \boldsymbol{P}\boldsymbol{x}(i), \quad i = 1, 2, \ldots, N$$

where $\boldsymbol{P}$ is the $n \times n$ diagonal matrix

$$\boldsymbol{P} = \operatorname{diag}\{a_1^{-1}, a_2^{-1}, \ldots, a_n^{-1}\}$$

We will use below the new data $\boldsymbol{y}(i)$, $i = 1, 2, \ldots, N$ instead of the $\boldsymbol{x}(i)$, $i = 1, 2, \ldots, N$.

The covariance matrix of the transformed data can be expressed as

$$\boldsymbol{R} = E\left[\boldsymbol{y}(i)\boldsymbol{y}^H(i)\right] = \boldsymbol{P}\boldsymbol{M}\boldsymbol{P}^H = \boldsymbol{R}_N + p_S \boldsymbol{e}\boldsymbol{e}^T \quad (2)$$

where the $n \times n$ diagonal matrix

$$\boldsymbol{R}_N = \operatorname{diag}\{p_1, p_2, \ldots, p_n\}$$

T denotes transpose. Here $p_k = \sigma_k^2/|a_k|^2$, $k = 1, 2, \ldots, n$ are unknown noise powers of the transformed data, $\boldsymbol{e} = (1, 1, \ldots, 1)^T$. Hence, the problem of estimation of the parameter p_S from the vector array data (1) can be reformulated as the same problem for the new vector data $\boldsymbol{y}(i)$ with the covariance matrix (2).

The log-likelihood function can be expressed as:

$$L = -\log\det \boldsymbol{R} - \operatorname{Tr}(\boldsymbol{R}^{-1}\hat{\boldsymbol{R}}) \quad (3)$$

where Tr denotes the trace of matrix and $\hat{\boldsymbol{R}}$ is the $n \times n$ sample covariance matrix

$$\hat{\boldsymbol{R}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{y}(i)\boldsymbol{y}^H(i) \quad (4)$$

The ML equations can be written as

$$\frac{\partial L}{\partial p_S} = 0, \qquad \frac{\partial L}{\partial p_k} = 0, \quad k = 1, 2, \ldots, n \quad (5)$$

for $p_S = \hat{p}_S$, $p_k = \hat{p}_k$, where $\hat{p}_S$ and $\hat{p}_k$ are the ML estimates of signal power and of kth noise power, respectively.

After straightforward calculations we find that the ML equations (5) can be expressed as

$$\left(1 + \hat{p}_S \sum_{l=1}^{n} \frac{1}{\hat{p}_l}\right) \sum_{l=1}^{n} \frac{1}{\hat{p}_l} = \sum_{l=1}^{n}\sum_{m=1}^{n} \frac{\hat{R}_{ml}}{\hat{p}_l \hat{p}_m},$$

$$\left(\hat{R}_{kk} - \hat{p}_k + \hat{p}_S\right)\left(1 + \hat{p}_S \sum_{l=1}^{n} \frac{1}{\hat{p}_l}\right) =$$

$$\hat{p}_S \sum_{l=1}^{n} \left(\frac{\hat{R}_{lk} + \hat{R}_{kl}}{\hat{p}_l}\right), \qquad k = 1, 2, \ldots, n$$

where $\hat{R}_{kl}$ is the (k, l)th element of sample covariance matrix (4).

Assume now that the SNR is low while the number of snapshots is large, i.e., $p_S \ll p_k$, $k = 1, 2, \ldots, n$, $N \gg 1$. After some straightforward algebra using these assumptions and the ML equations, we get the following simple AML estimator of signal power (see [5] for details):

$$\hat{p}_S = \sum_{\substack{k,l=1 \\ k \neq l}}^{n} \frac{\hat{R}_{kl}}{\hat{R}_{kk}\hat{R}_{ll}} \Bigg/ \sum_{\substack{k,l=1 \\ k \neq l}}^{n} \frac{1}{\hat{R}_{kk}\hat{R}_{ll}} \quad (6)$$

The estimator (6) is expressed in a form conductive to simple implementation. This formulation is devoid of matrix inversion or eigendecomposition. It should be also noted that (6) can be considered as a variant of matched processing for arbitrary and unknown sensor noise powers case. Therefore, it can be extended to the case of well separated multiple signals with unknown directions of arrival. In this case one should estimate the function $p_S(\theta)$ within angular area of interest (if the signals are assumed to be plane waves). The steering towards the chosen set of angles can be done using matrix $\boldsymbol{P}$.

3. CRAMER–RAO BOUND

It is well known that for any unbiased estimate of vector parameter $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)^T$ (m is the total number of unknowns), the CRB is given by the diagonal elements of inverted Fisher information matrix $\boldsymbol{J}^{-1}$: $\mathrm{CRB}(\hat{\theta}_i) = [\boldsymbol{J}^{-1}]_{ii}$. In turn, the (i, k)th element of Fisher information matrix for multivariate Gaussian complex vector data with zero mean can be expressed by well known Bangs' formula

$$J_{ik} = N \operatorname{Tr}\left(\boldsymbol{R}^{-1}\frac{\partial \boldsymbol{R}}{\partial \theta_i}\boldsymbol{R}^{-1}\frac{\partial \boldsymbol{R}}{\partial \theta_k}\right)$$

It can be shown using this equation that (see [6] for details):

$$\mathrm{CRB}(\hat{p}_S) = \frac{1}{N}\frac{D(D+p_S^2 G)}{(D-1)^2/p_S^2+(D-2)G},$$

$$D = 1 + p_S\sum_{k=1}^{n}\frac{1}{p_k}, \quad G = \sum_{k=1}^{n}\frac{1}{p_k^2 D - 2p_S p_k} \tag{7}$$

4. STATISTICAL PERFORMANCE

It can be shown that (6) is the *asymptotically unbiased* estimator, i.e.,

$$E\left[\hat{p}_S\right] = p_S + o\left(\frac{1}{\sqrt{N}}\right)$$

The asymptotic variance of the estimator (6) can be expressed as

$$\mathrm{var}(\hat{p}_S) = \frac{1}{N}\left[p_S^2 + \left[2p_S\sum_{l=1}^{n}\frac{p_l}{(p_l+p_S)^2}\left(\sum_{\substack{k=1\\k\neq l}}^{n}\frac{1}{p_k+p_S}\right)^2 + \sum_{\substack{k,l=1\\k\neq l}}^{n}\frac{p_k p_l}{(p_k+p_S)^2(p_l+p_S)^2}\right]\Bigg/ \left(\sum_{\substack{k,l=1\\k\neq l}}^{n}\frac{1}{(p_k+p_S)(p_l+p_S)}\right)^2\right] + o\left(\frac{1}{N}\right)$$

Comparison of the asymptotic variance of the estimator (6) with the exact CRB (7) in practically important cases gives the following results.

4.1. Low SNR Case

In this case we have

$$\lim_{p_S\to 0}\mathrm{CRB}(\hat{p}_S) = \frac{1}{N}\left(\sum_{\substack{k,l=1\\k\neq l}}^{n}\frac{1}{p_k p_l}\right)^{-1} =$$

$$\lim_{p_S\to 0}\mathrm{var}(\hat{p}_S) + o\left(\frac{1}{N}\right)$$

Therefore, the AML estimator (6) approaches asymptotic efficiency as the SNR approaches zero.

4.2. High SNR Case

For high SNR, comparison of CRB and asymptotic variance yields

$$\lim_{p_S\to\infty}\mathrm{CRB}(\hat{p}_S) = \frac{1}{N}p_S^2 =$$

$$\lim_{p_S\to\infty}\mathrm{var}(\hat{p}_S) + o\left(\frac{1}{N}\right)$$

It implies that although the derivation of AML estimator (6) was carried out under the low SNR assumption, it also approaches asymptotic efficiency for high SNR.

4.3. Two Sensor Case

For $n = 2$ the AML estimator (6) becomes the ML one for any SNR value and any number of samples N: $\hat{p}_S = (\hat{R}_{12} + \hat{R}_{21})/2$. Its variance and the CRB also coincide exactly.

4.4. Identical Noise Powers Case

Assume that the noise has identical powers in each sensor, i.e., $p_1 = p_2 = \cdots = p_n$. Designating $p_k = p_N$, $k = 1, 2, \ldots, n$, we have

$$\mathrm{CRB}(\hat{p}_S) = \frac{1}{N}\left(p_S^2 + \frac{2p_S p_N}{n} + \frac{p_N^2}{n(n-1)}\right) =$$

$$\mathrm{var}(\hat{p}_S) + o\left(\frac{1}{N}\right)$$

The last equation implies that (6) is asymptotically efficient estimator in the case of identical noise powers and arbitrary SNR.

4.5. Case of Several Distinguished Sensors

Assume now that the noise powers in arbitrary K sensors ($K < n-1$) with the numbers $l_1, \ldots, l_K$ from the total number of array sensors are much higher than that in the other $n-K$ sensors, i.e., $p_{l_m} \gg p_k$, $m = 1, 2, \ldots, K$, $k \neq l_1, l_2, \ldots, l_K$. Assume also that the noise powers in the other $n-K$ sensors are identical, i.e., $p_k = p_N$, $k \neq l_1, l_2, \ldots, l_K$. It can be shown that

$$\lim_{p_{l_1},\ldots,p_{l_K}\to\infty}\mathrm{CRB}(\hat{p}_S) =$$

$$\frac{1}{N}\left(p_S^2 + 2\frac{p_S p_N}{(n-K)} + \frac{p_N^2}{(n-K)(n-K-1)}\right) =$$

$$\lim_{p_{l_1},\ldots,p_{l_K}\to\infty}\mathrm{var}(\hat{p}_S) + o\left(\frac{1}{N}\right)$$

As a result, the distinguished sensors with powerful noises have no influence on the CRB and on the variance of the AML estimator (6). Indeed, the results of the noisy sensors are included with the low weights in the AML estimator (6). The estimator (6) retains the asymptotic efficiency in the case considered.

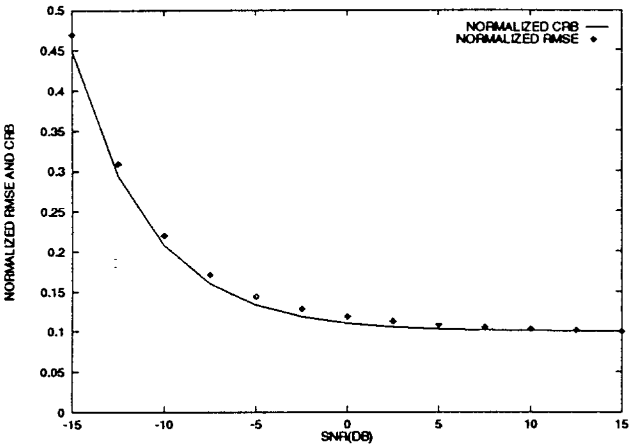

Figure 1: Normalized experimental RMSE and normalized CRB versus SNR for the first noise scenario.

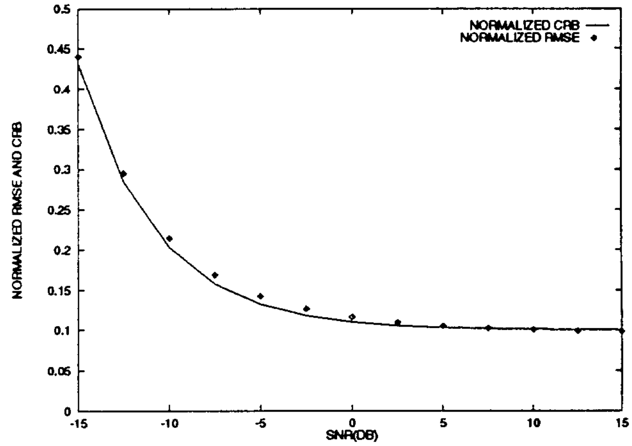

Figure 2: Normalized experimental RMSE and normalized CRB versus SNR for the second noise scenario.

5. SIMULATION RESULTS

We examine the performance of the derived AML estimator (6) by comparison the root-mean-square error (RMSE) of signal power estimation and the CRB.

We assumed the uniform linear array of ten sensors and assumed also that the noise was uncorrelated from sensor to sensor and uncorrelated with the signal too. A single signal was assumed to impinge on the array from the normal direction. The number of samples taken was 100 and a total of 100 independent simulation runs have been performed to compute the RMSE for each SNR value. The SNR was defined as

$$\mathrm{SNR} = (p_S/n)\sum_{k=1}^{n}(1/p_k)$$

In the first example, we assumed the following sensor noise scenario: $p_1 = 1.00$, $p_2 = 4.57$, $p_3 = 3.13$, $p_4 = 7.89$, $p_5 = 18.01$, $p_6 = 0.57$, $p_7 = 1.54$, $p_8 = 5.13$, $p_9 = 2.77$, and $p_{10} = 12.39$.

In the second example, we assumed the scenario: $p_1 = 0.67$, $p_2 = 1.23$, $p_3 = 0.86$, $p_4 = 1.54$, $p_5 = 1.61$, $p_6 = 1.04$, $p_7 = 0.92$, $p_8 = 0.58$, $p_9 = 0.89$, and $p_{10} = 1.11$.

Figures 1 and 2 show the resulting normalized experimental RMSE of estimator (6) and normalized CRB versus SNR for the first and second example, respectively. The normalized RMSE and CRB are here defined as:

$$\text{normalized RMSE} = \frac{\mathrm{RMSE}}{p_S}$$

$$\text{normalized CRB} = \frac{\sqrt{\mathrm{CRB}}}{p_S}$$

The results of simulations verify that the RMSE of the AML estimator is very close to CRB for a wide SNR range and, in fact, has the same asymptotic performance, as the exact (optimal) ML estimator.

REFERENCES

1. Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. ASSP*, **34**, pp. 1081–1089, Oct. 1986.
2. J.F. Böhme, "Source parameter estimation by approximate maximum likelihood and nonlinear regression,"*IEEE J. Oceanic Eng.*, **OE-10**, pp. 206–212, 1985.
3. B.D. Van Veen and L.L Sharf, "Estimation of structured covariance matrices and multiple window analysis," *IEEE Trans. ASSP*, **38**, pp. 1467–1472, Aug. 1990.
4. A.B. Gershman, V.I. Turchin and R.A. Ugrinovsky, "Simple maximum-likelihood estimator of structured covariance parameters," *Electron. Lett.*, **28**, no. 18, pp.1677-1678, Aug. 1992.
5. A.B. Gershman, A.L. Matveyev and J.F. Böhme, "Maximum likelihood estimation of signal power in sensor array in the presence of unknown noise field," submitted to *IEE Proc. Radar, Sonar and Navig.*
6. A.B. Gershman and A.L. Matveyev, "Exact Cramer-Rao bound on signal power estimation in the presence of signal noises with nonidentical and unknown powers," submitted to *IEEE Trans. on Signal Processing.*