16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

DETERMINISTIC ML ESTIMATION FOR UNKNOWN NUMBERS OF SIGNALS

Pei-Jung Chung[1], *Christoph F. Mecklenbräuker*[2]

[1]School of Engineering and Electronics
The University of Edinburgh, UK
[2]Inst. of Communications and RF Engineering
Vienna University of Technology, Austria

p.chung@ed.ac.uk, cfm@nt.tuwien.ac.at

ABSTRACT

The knowledge about the number of signals plays a crucial role in array processing. The performance of most direction finding algorithms relies strongly on a correctly specified number of signals. When the number of signals is unknown, conventional approaches apply information theoretic criteria or multiple tests to estimate the number of signals and parameters of interest simultaneously. These methods usually compute ML estimates for a hierarchy of nested models. The total computational complexity is significantly higher than the standard ML procedure. In this contribution, we develop a novel ML approach that computes ML estimates only for the maximal hypothesized number of signals. Furthermore, we introduce a multiple hypothesis test to identify relevant components that are associated with the true DOA parameters. Numerical experiments show that the proposed method provides comparable estimation accuracy as the standard ML method does.

1. INTRODUCTION

The problem of estimating direction of arrival (DOA) is a key issue in array processing. Among existing methods, the maximum likelihood (ML) approach has the best statistical properties and is robust against small sample numbers, signal coherence and closely located sources.

The standard ML method assumes the number of signals, m, to be known and maximizes the likelihood function over an m-dimensional parameter space. When the number of signals is unknown, conventional approaches estimate DOA parameters along with m using the information theoretic criterion based methods [9, 11] or the multiple hypothesis tests [6, 7] in a sequential manner. Given the maximal hypothesized number of signals, M, these methods compute ML estimates for a hierarchy of nested models and select the one that best fits the underlying criterion. Since the likelihood function is optimized over a series of candidate models, the total computational cost can be significantly higher than standard ML estimation.

P.-J. Chung acknowledges support of her position from the Scottish Funding Council and their support of the Joint Research Institute with the Heriot-Watt University as a component part of the Edinburgh Research Partnership.

In this work, we suggest a novel procedure that *only* computes ML estimates for the maximal hypothesized model. The proposed approach is motivated by the fact that the ML estimator derived from an overestimated model order contains components that coincide with the true parameters [2]. We also introduce a multiple hypothesis test to identify relevant components. In the previous contribution based on stochastic signal models [3], the relevant components are chosen by thresholding the likelihood function. Here we consider deterministic signal models and select relevant components in a statistically justified manner. Since the multiple test takes data distribution into account, the resulting estimates should be more reliable than those of [3].

Although the primary purpose of the proposed algorithm is estimating DOA parameters for unknown numbers of signals, we can determine the number of signals based on the number of relevant components. Clearly, the proposed approach requires no sequential maximization over various parameter spaces and is computationally more attractive than conventional methods.

In the following, we give a brief description of deterministic signal models. Section 3 includes asymptotic properties of ML estimation of misspecified numbers of signals. In Section 4, we develop the proposed ML estimation algorithm for unknown numbers of signals. Simulation results are presented in Section 5. Concluding remarks are given in Section 6.

2. PROBLEM FORMULATION

Consider an array of n sensors receiving m narrow band signals emitted by far-field sources located at $\boldsymbol{\theta}_m = [\,\theta_1,\ldots,\theta_m]^T$. The array output $\boldsymbol{x}(t)$ is described as

$$\boldsymbol{x}(t) = \boldsymbol{H}_m(\boldsymbol{\theta}_m)\boldsymbol{s}_m(t) + \boldsymbol{n}(t), \quad t = 1,\ldots,T, \qquad (1)$$

where the ith column $\boldsymbol{d}(\theta_i)$ of the matrix

$$\boldsymbol{H}_m(\boldsymbol{\theta}_m) = [\boldsymbol{d}(\theta_1) \cdots \boldsymbol{d}(\theta_i) \cdots \boldsymbol{d}(\theta_m)] \tag{2}$$

represents the steering vector associated with the signal arriving from θ_i. The $n \times m$ matrix $\boldsymbol{H}_m(\boldsymbol{\theta}_m)$ is assumed to be full rank. The complex signal waveform $\boldsymbol{s}_m(t) = [s_1(t), \ldots, s_m(t)]^T$ is considered as deterministic and unknown. The noise vector $\boldsymbol{n}(t)$ is independent, identically complex normally distributed with zero mean and covariance matrix $\nu \boldsymbol{I}_n$, where ν denotes the unknown noise spectral parameter.

Based on the observations $\{\boldsymbol{x}(t)\}_{t=1}^T$ and a pre-specified number of signals m, the ML estimate is obtained by minimizing the negative concentrated likelihood function [1]

$$\begin{aligned} \hat{\boldsymbol{\theta}}_m &= \arg\min_{\boldsymbol{\theta}_m} l_T(\boldsymbol{\theta}_m), \\ l_T(\boldsymbol{\theta}_m) &= \mathrm{tr}[\boldsymbol{P}_m^{\perp}(\boldsymbol{\theta}_m)\hat{\boldsymbol{C}}_x\,], \end{aligned} \tag{3}$$

where $\boldsymbol{P}_m^{\perp}(\boldsymbol{\theta}_m) = \boldsymbol{I} - \boldsymbol{P}_m(\boldsymbol{\theta}_m)$ and $\boldsymbol{P}_m(\boldsymbol{\theta}_m)$ is the projection matrix into the column space of $\boldsymbol{H}_m(\boldsymbol{\theta}_m)$. The sample covariance matrix $\hat{\boldsymbol{C}}_x = \frac{1}{T}\sum_{t=1}^T \boldsymbol{x}(t)\boldsymbol{x}(t)^H$.

The problem of central interest is to estimate the DOA parameters when the true number of signals, m_0, is unknown. In the following we shall develop an algorithm that requires only an upper bound on the number of signals. The proposed approach is motivated by the fact that for an overestimated number of signals, the ML estimate contains components that coincide with the true parameters.

3. ML ESTIMATION FOR MISSPECIFIED NUMBERS OF SIGNALS

It is well established that the ML estimator converges to the true parameter $\boldsymbol{\theta}_0$ with increasing sample size [5, 8] when the number of signals, m, is correctly specified. For $m \neq m_0$, we applied the general theory of misspecified nonlinear least regression models [10] to study the asymptotic behavior of the ML estimator. In [2, 4], we showed that for a misspecified m, the ML estimator $\hat{\boldsymbol{\theta}}_m$ converges to a well defined limit $\boldsymbol{\theta}_m^*$ that minimizes the ensemble average of the concentrated likelihood function

$$l(\boldsymbol{\theta}_m) = E\left\{\mathrm{tr}[\boldsymbol{P}_m^{\perp}(\boldsymbol{\theta}_m)\boldsymbol{x}(t)\boldsymbol{x}(t)^H]\right\}. \tag{4}$$

The signal subspace computed at $\boldsymbol{\theta}_m^*$ is related to the true signal subspace $\mathrm{sp}(\boldsymbol{H}_{m_0}(\boldsymbol{\theta}_0))$ with as follows.

Consistency Property

(a) For $m < m_0$,

$$\mathrm{sp}(\boldsymbol{H}_m(\boldsymbol{\theta}_m^*)) \subset \mathrm{sp}(\boldsymbol{H}_{m_0}(\boldsymbol{\theta}_0)) \tag{5}$$

and the elements of $\boldsymbol{\theta}_m^*$ coincide with m elements of $\boldsymbol{\theta}_0$ for widely separated sources.

(b) For $m > m_0$,

$$\mathrm{sp}(\boldsymbol{H}_m(\boldsymbol{\theta}_m^*)) \supset \mathrm{sp}(\boldsymbol{H}_{m_0}(\boldsymbol{\theta}_0)) \tag{6}$$

and $\boldsymbol{\theta}_m^*$ contains m_0 components equal to those of $\boldsymbol{\theta}_0$. The remaining $(m - m_0)$ components of are unpredictable.

Proof Details can be found in [2]. □

Part (b) suggests that when exact knowledge about the number of signals is not available, the estimates for the true parameter $\boldsymbol{\theta}_0$ can be obtained by computing the ML estimates for an overparameterized model with $m > m_0$.

4. ROBUST ML ESTIMATION

Let M denote an upper bound on the number of signals. Motivated by the consistency property discussed previously, we propose to compute the ML estimate for the overparameterized model $m = M$ and select relevant components that are associated with the true parameters. More specifically, the proposed algorithm computes the ML estimate $\hat{\boldsymbol{\theta}}_M$ by minimizing the negative log-likelihood function (3) over an M-dimensional space

$$\hat{\boldsymbol{\theta}}_M = \arg\min_{\boldsymbol{\theta}_M} l_T(\boldsymbol{\theta}_M). \tag{7}$$

Since $M \geq m_0$, the $M \times 1$ vector $\hat{\boldsymbol{\theta}}_M = [\hat{\theta}_1, \ldots, \hat{\theta}_M]^T$ contains more elements than the $m_0 \times 1$ true parameter vector $\boldsymbol{\theta}_0$ does. As discussed previously, for large T, a subset of the elements in $\hat{\boldsymbol{\theta}}_M$ may coincide with those of $\boldsymbol{\theta}_0$. The elements of $\hat{\boldsymbol{\theta}}_M$ that are associated with those of $\boldsymbol{\theta}_0$, are referred to as *relevant* components. The remaining $(M - m_0)$ components of $\hat{\boldsymbol{\theta}}_M$ are referred to as *redundant* components.

A key step in such max-search procedure is identification of relevant components. In [3], the relevant components are selected by thresholding the likelihood function because the redundant components do not change the value of the likelihood function. The threshold used in [3] is chosen in an *ad hoc* manner. In this contribution, we consider the following hypothesis test to validate the ith component:

$$\begin{aligned} H_i &: \boldsymbol{x}(t) = \boldsymbol{H}_{M-1}(\tilde{\boldsymbol{\theta}}_i)\tilde{\boldsymbol{s}}_{M-1}(t) + \boldsymbol{n}(t) \\ A_i &: \boldsymbol{x}(t) = \boldsymbol{H}_M(\hat{\boldsymbol{\theta}}_M)\boldsymbol{s}_M(t) + \boldsymbol{n}(t) \end{aligned} \tag{8}$$

where H_i and A_i represent the null hypothesis and the alternative, respectively. The $(M-1) \times 1$ vector

$$\tilde{\boldsymbol{\theta}}_i = [\hat{\theta}_1 \cdots \hat{\theta}_{i-1}\ \hat{\theta}_{i+1} \cdots \hat{\theta}_M]^T \tag{9}$$

contains all elements of $\hat{\boldsymbol{\theta}}_M$ except the ith component. The $n \times (M-1)$ matrix $\boldsymbol{H}_{M-1}(\tilde{\boldsymbol{\theta}}_i)$ contains steering vectors corresponding to the DOA parameters in $\hat{\boldsymbol{\theta}}_{M-1}$.

Applying the likelihood ratio principle, we obtain the following test statistic for testing H_i against A_i:

$$T_i = \log\left(\frac{\mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_{M-1}(\tilde{\boldsymbol{\theta}}_i))\hat{\boldsymbol{C}}_{\boldsymbol{x}}]}{\mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_M(\hat{\boldsymbol{\theta}}_M))\hat{\boldsymbol{C}}_{\boldsymbol{x}}]}\right) \tag{10}$$

$$= \log\left(1 + \frac{n_1}{n_2}F_i\right). \tag{11}$$

Under the null hypothesis H_i, the statistic

$$F_i = \frac{n_2}{n_1}\frac{\mathrm{tr}[(\boldsymbol{P}_M(\hat{\boldsymbol{\theta}}_M) - \boldsymbol{P}_{M-1}(\tilde{\boldsymbol{\theta}}_i))\hat{\boldsymbol{C}}_{\boldsymbol{x}}]}{\mathrm{tr}[(\boldsymbol{I} - \boldsymbol{P}_M(\hat{\boldsymbol{\theta}}_M))\hat{\boldsymbol{C}}_{\boldsymbol{x}}]} \tag{12}$$

is F_{n_1,n_2}–distributed with the degrees of freedom n_1, n_2 given by [6]

$$n_1 = 3T, \quad n_2 = T(2n - 2M - 1). \tag{13}$$

The component $\hat{\theta}_i$ is considered as relevant if

$$F_i \geq t_\alpha, \tag{14}$$

where t_α is the threshold for a given significance level α.

Compared to the thresholding strategy proposed in [3], the hypothesis test (8) incorporate data distribution into the selection step. Hence the probability of correct detection is expected to be higher than that of [3]. Using (14) we control the significance level of individual tests H_i, $(i = 1, \ldots, M)$ rather than the global test level. A more sophisticated test procedure taking the multiplicity into account is still under investigation.

Let the relevant vector $\hat{\boldsymbol{\theta}}_0 = [\hat{\theta}_{(1)}, \ldots, \hat{\theta}_{(k)}]$ include all the relevant components validated by the test (8). Although the primary purpose of the proposed algorithm is to estimate wave parameters, the number of relevant components provides indirectly an estimate for the number of signals.

In summary, given an upper bound on the number of the signals M and the threshold t_α, the proposed algorithm proceeds as follows.

Input: $\{\boldsymbol{x}(t) : t = 1, \ldots, T\}$, M, t_α.

1. Find the ML estimate $\hat{\boldsymbol{\theta}}_M$ by (7).
2. Compute the statistic F_i, $i = 1, \ldots, M$, by (12).
3. Test H_i against A_i to select relevant components.

   $\hat{\theta}_i$ is relevant if $F_i \geq t_\alpha$.

Output: $\hat{\boldsymbol{\theta}}_0 = [\hat{\theta}_{(1)}, \hat{\theta}_{(2)}, \ldots, \hat{\theta}_{(k)}]^T$.

Table 1: Deterministic ML Estimation Algorithm for Unknown Numbers of Signals.

5. SIMULATION

We test the proposed algorithm on simulated data. In particular, we shall investigate whether the relevant estimates provide useful information about the true parameters.

In the numerical experiments, a uniform linear of 10 sensors with inter-element spacings of half a wavelength is employed. The narrow band signals are generated by $m_0 = 2$ uncorrelated signals located at $[\,26°\ 36°\,]$ of various strengths. The two signals are separated about half of the beamwidth. The difference of signal strengths is $[\,1 \quad 0\,]$ dB where 0 dB corresponds to the reference signal. The signal to noise ratio (SNR), defined as $10\log\left(E(|s_i(t)|^2)/\nu\right)$ for the ith signal, varies from -10 to 10 dB in a 2 dB step. We generate $T = 100$ snapshots for each of the 200 trials performed. The upper bound on the number of signals is chosen to be $M = 3,\ 4$. The significance level $\alpha = 0.1$. For comparison, we apply the ML approach to the same batch of data using the correct number of signals, $m_0 = 2$.

Fig. 1 shows the bias of the relevant components $\hat{\theta}_{(1)}$, $\hat{\theta}_{(2)}$, respectively. For all three curves corresponding to $M = 2, 3, 4$ the bias is less than 0.1 degree over the entire SNR range. Since the results are obtained from finite samples, we conjecture that $\hat{\theta}_{(1)}$ and $\hat{\theta}_{(2)}$ are asymptotically bias free.

The standard deviation (square root of empirical variance) of $\hat{\theta}_{(1)}$ and $\hat{\theta}_{(2)}$ are presented in Fig. 2. All three curves decline with increasing SNR. The estimates obtained from the correct number of signals $M = m_0 = 2$ lead to the smallest variances. In the upper panel we can observe that the first relevant component $\hat{\theta}_{(1)}$ is better estimated by $M = 4$ than $M = 3$. However, in the lower panel, $M = 4$ results in a larger variance than $M = 3$. This implies that the variance of each component does not necessarily increase with a larger degree of mismatch.

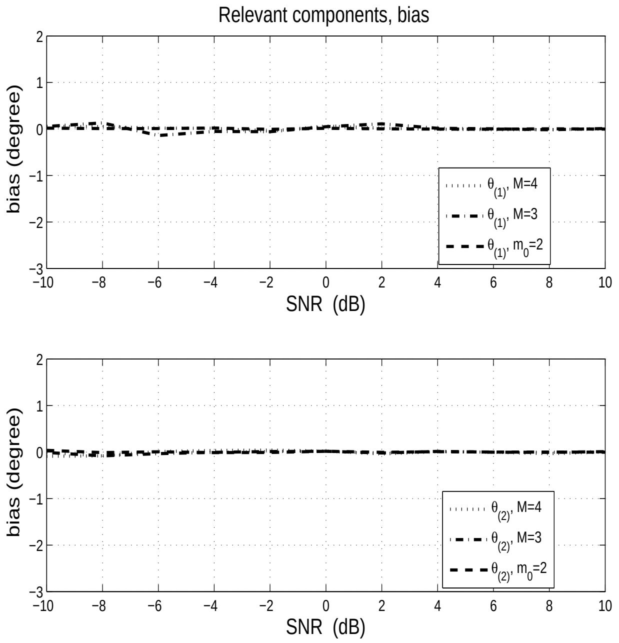

Fig. 1. Bias. $M = 2, 3, 4$. The true DOA parameter $\boldsymbol{\theta}_0 = [26° \; 36°]$, SNR $= [-10:2:10]$ dB.

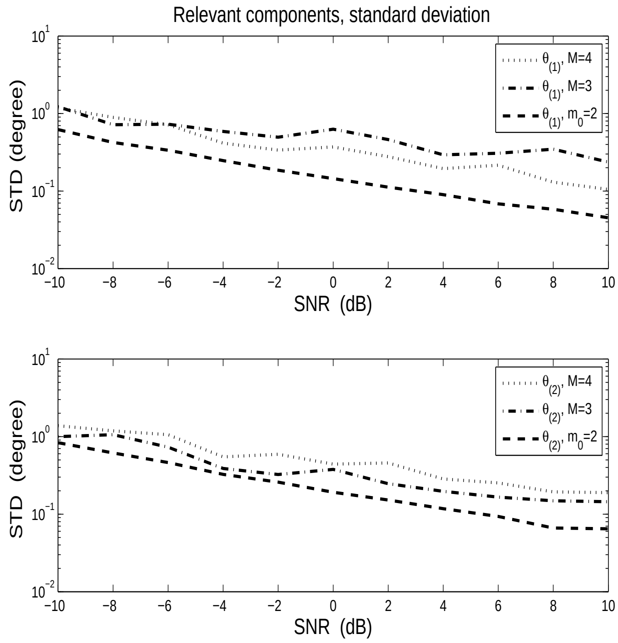

Fig. 2. Standard deviation of $\hat{\theta}_{(i)}, \; i = 1, 2$.

In Fig. 3, we compare the probability of correct detection of the proposed algorithm with the popular MDL criterion based approach [9]. By "correct detection", we mean that the number of relevant components equals the true number of signals. For low SNR region, -10 to -8 dB, both $M = 3$ and $M = 4$ have a higher probability of correct detection. For SNR -6 to 10 dB, MDL achieves 100% probability of correct detection, while $M = 3, 4$ increase from 90% to over 96%. The curve associated with $M = 3$ shows a slightly higher probability of correct detection than that of $M = 4$. The results suggest that without computing ML estimates for all candidate models, the proposed algorithm has comparable performance as the computationally more involved MDL approach.

6. CONCLUSION

We developed a ML estimation procedure for unknown numbers of signals based on deterministic signal models. The suggested algorithm computes ML estimates only for the maximal hypothesized number of signals. Compared to traditional methods for joint parameter estimation and signal detection, the proposed max-search approach avoids the full search process through a series of nested models and leads to significant improvement in computational efficiency. The relevant components that are associated with the true DOA parameters are identified by a multiple hypothesis test. Numerical results showed that the proposed algorithm achieves comparable estimation accuracy as the standard ML approach does. Furthermore, the number of signals can be accurately determined by the number of relevant components. The proposed algorithm provides a computationally attractive alternative to existing joint parameter estimation and signal detection methods.

7. REFERENCES

[1] J. F. Böhme. Statistical array signal processing of measured sonar and seismic data. In *Proc. SPIE 2563 Advanced Signal Processing Algorithms*, pages 2–20, San Diego, Jul 1995.

[2] P.-J. Chung. ML estimation under misspecified number of signals. In *Proc. of the 39th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Oct 30 - Nov 2, 2005.

[3] P.-J. Chung. Robust ML estimation for unknown numbers of signals. In *Proc. EUSIPCO*, Poznań, Poland, September 2007.

[4] Pei-Jung Chung. Performance analysis of ML estimation under misspecified numbers of signals. In *IEEE*

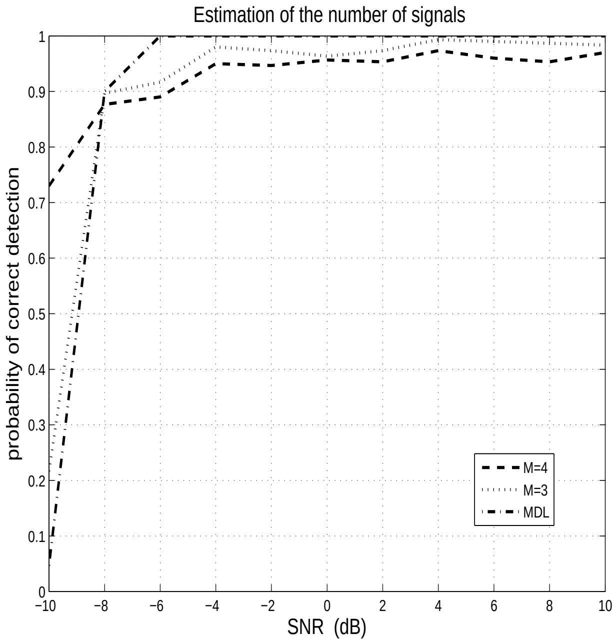

Fig. 3. Probability of correct detection.

Proc. Sensor Array and Multichannel Signal Processing Workshop, Boston, USA, July 2006.

[5] D. Kraus. *Approximative Maximum-Likelihood-Schätzung und verwandte Verfahren zur Ortung und Signalschätzung mit Sensorgruppen*. Dr.–Ing. Dissertation, Faculty of Electrical Engineering, Ruhr–Universität Bochum, Shaker Verlag, Aachen, 1993.

[6] D. Maiwald. *Breitbandverfahren zur Signalentdeckung und –ortung mit Sensorgruppen in Seismik– und Sonaranwendungen*. Dr.–Ing. Dissertation, Dept. of Electrical Engineering, Ruhr–Universität Bochum, Shaker Verlag, Aachen, 1995.

[7] P.-J.Chung, J. F. Böhme, C. F. Mecklenbräuker, and A. O. Hero. Detection of the number of signals using the Benjamini-Hochberg procedure. *IEEE Trans. Signal Processing*, 55(6):2497–2508, June 2007.

[8] P. Stoica and A. Nehorai. Performance study of conditional and unconditional direction–of–arrival estimation. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 38(10):1783–1795, October 1990.

[9] M. Wax and I. Ziskind. Detection of the number of coherent signals by the MDL principle. *IEEE ASSP*, 37(8):1190–1196, August 1989.

[10] Halbert White. Consequences and detection of misspecified nonlinear regression models. *Journal of American Statistical Association*, 76(374):419–433, 1981.

[11] M. Wong, Q. T. Zou, and J. P. Reilly. On information theoretic criterion for determining the number of signals in high resolution array processing. *IEEE Trans. Acoust., Speech, Signal Processing*, 38(11):1959–1971, November 1990.