

MODIFIED IQML AND A STATISTICALLY EFFICIENT METHOD FOR DIRECTION ESTIMATION WITHOUT EIGENDECOMPOSITION

Martin Kristensson Magnus Jansson Björn Ottersten

Dept. of Signals, Sensors & Systems, Royal Institute of Technology, SE-100 44 Stockholm, Sweden

ABSTRACT

This paper deals with direction estimation of signals impinging on a uniform linear sensor array. A well known algorithm for this problem is IQML. Unfortunately, the IQML estimates are in general biased, especially in noisy scenarios. We propose a modification of IQML (MIQML) that gives consistent estimates at approximately the same computational cost. In addition, an algorithm with an estimation error covariance which is asymptotically identical to the asymptotic Cramér-Rao lower bound is presented. The optimal algorithm resembles weighted subspace fitting or MODE, but achieves optimal performance without having to compute an eigendecomposition of the sample covariance matrix.

1. INTRODUCTION

The problem of estimating directions of arrival using sensor arrays has been intensively investigated [3]. The well known IQML (Iterative Quadratic Maximum Likelihood) algorithm is closely connected with the deterministic maximum likelihood estimator (DML) and also shares the same properties in low noise cases. In addition, the IQML algorithm is attractive also from a computational point of view.

However, as noted in [6], the IQML algorithm does not converge to the DML estimates. In fact, the IQML estimates are biased in the general case. That is, even if IQML operates on the true covariance matrix, it does not deliver the true directions. This biased behavior of IQML is of course a limitation.

In this contribution, we first propose an IQML like estimator that in contrast to IQML is consistent and at the same time shares the attractive computational properties of the IQML approach. Hereafter, we will refer to this modified IQML estimator as MIQML. We also propose a statistically efficient extension of MIQML which we call WSF-E (Weighted Subspace Fitting without Eigendecomposition). WSF-E is closely connected to the well known weighted subspace fitting estimator (WSF) [8].

The paper is organized as follows. In the next two sections the data model and some notations are introduced. We also comment on the importance of using correct constraints

on the parameters in the model. In Section 4, the IQML algorithm and its properties are discussed followed by the first new algorithm in Section 5. Thereafter, the optimal algorithm is presented followed by some simulation results. Finally, some conclusions are given.

2. MODEL DESCRIPTION

In this paper, we assume that the output vector $\mathbf{x}(t) \in \mathcal{C}^{m \times 1}$ of a uniform linear array is given by the model

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t), \quad (1)$$

where $\mathbf{s}(t) \in \mathcal{C}^{d \times 1}$ contains the d narrowband transmitted signals and the circular zero mean Gaussian noise process $\mathbf{n}(t) \in \mathcal{C}^{m \times 1}$ is spatially and temporally white. $\mathbf{A}(\boldsymbol{\theta})$ is the $m \times d$ array steering matrix whose k^{th} column is $\mathbf{a}(\theta_k) = [1 \quad e^{j2\pi\Delta\sin(\theta_k)} \quad e^{j(m-1)2\pi\Delta\sin(\theta_k)}]^T$, where $(\cdot)^T$ denotes transpose, θ_k is the k^{th} direction of arrival and Δ is the element spacing measured in wavelengths. We will consider the signals $\mathbf{s}(t)$ as either being unknown deterministic vectors or as being realizations of a circular Gaussian process. In the statistical analysis we use the stochastic assumption on $\mathbf{s}(t)$ and let $\mathbf{S} = E\{\mathbf{s}(t)\mathbf{s}^*(t)\}$ denote the covariance of $\mathbf{s}(t)$. Here $(\cdot)^*$ denotes complex conjugate transpose. The array output is assumed to be white in time with covariance

$$\mathbf{R} = E\{\mathbf{x}(t)\mathbf{x}^*(t)\} = \mathbf{A}(\boldsymbol{\theta})\mathbf{S}\mathbf{A}^*(\boldsymbol{\theta}) + \sigma^2\mathbf{I}. \quad (2)$$

An estimate of \mathbf{R} is given by $\hat{\mathbf{R}} = N^{-1} \sum_{t=1}^N \mathbf{x}(t)\mathbf{x}^*(t)$, where N is the number of array snapshots available.

3. NOISE SUBSPACE PARAMETERIZATION

It is well known that it is possible to find a linear parameterization of the null space of $\mathbf{A}(\boldsymbol{\theta})$, see e.g., [1, 7]. That is, it is possible to find a full rank matrix $\mathbf{B}(\mathbf{b})$ such that $\mathbf{B}^*(\mathbf{b})\mathbf{A}(\boldsymbol{\theta}) = \mathbf{0}$. This $m \times (m-d)$ matrix is given by

$$\mathbf{B}^T(\mathbf{b}) = \begin{bmatrix} b_0 & \dots & b_d & 0 & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & b_0 & \dots & b_d \end{bmatrix}, \quad (3)$$

where the parameter vector $\mathbf{b} = [b_0, \dots, b_d]^T$ is defined by

$$b_0 + b_1 z^1 + \dots + b_d z^d = b_d \prod_{k=1}^d (z - e^{-j 2\pi \Delta \sin(\theta_k)}).$$

For exact equivalence of the two parameterizations $\mathbf{A}(\boldsymbol{\theta})$ and $\mathbf{B}(\mathbf{b})$, it is necessary to restrict the parameter \mathbf{b} to only include polynomials with roots on the unit circle. However, in this paper, we consider the slightly weaker condition that the polynomial coefficients are complex conjugate symmetric, i.e., $b_k = b_{d-k}^c$, $k = 0, 1, \dots, d$, where $(\cdot)^c$ denotes complex conjugate. This symmetry can be exploited in the parameterization if we introduce the real valued $d + 1$ parameter vector $\boldsymbol{\mu}$ and the matrix \mathbf{K} with ones on the anti-diagonal and zeros elsewhere. Now, for d odd the complex parameter vector \mathbf{b} can be expressed in the following way

$$\mathbf{b} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I} & j\mathbf{I} \\ \mathbf{K} & -j\mathbf{K} \end{bmatrix} \boldsymbol{\mu} \triangleq \Phi \boldsymbol{\mu} \quad (4)$$

and for d even

$$\mathbf{b} = \frac{1}{\sqrt{2}} \begin{bmatrix} \mathbf{I} & \mathbf{0} & j\mathbf{I} \\ \mathbf{0} & \sqrt{2} & \mathbf{0} \\ \mathbf{K} & \mathbf{0} & -j\mathbf{K} \end{bmatrix} \boldsymbol{\mu} \triangleq \Phi \boldsymbol{\mu}. \quad (5)$$

Note the unitary mapping between \mathbf{b} and $\boldsymbol{\mu}$ ($\Phi^* \Phi = \mathbf{I}$). There are $d + 1$ real valued parameters in $\boldsymbol{\mu}$ compared to the d directions in $\boldsymbol{\theta}$. Therefore, it is necessary to use either a norm or linear constraint on the parameters [7, 5, 4]. As will be shown later, the norm constraint and the linear constraint in general behave differently and the choice of constraint even affects the consistency of the direction estimates. The modified IQML algorithm to be presented strongly relies on the norm constraint and the unitary mapping in Equations (4) and (5). For a more complete discussion on norm and linear constraints, see [4].

4. IQML

Using a deterministic model of the transmitted signals, the maximum likelihood estimate of the directions of arrival can be calculated from the minimizing argument of

$$V_{\text{DML}}(\boldsymbol{\mu}) = \text{Tr}\{\mathbf{P}_A^\perp \hat{\mathbf{R}}\} = \text{Tr}\{\mathbf{P}_B \hat{\mathbf{R}}\},$$

where $\mathbf{P}_A^\perp = \mathbf{I} - \mathbf{A}(\mathbf{A}^* \mathbf{A})^{-1} \mathbf{A}^*$ and $\mathbf{P}_B = \mathbf{B}(\mathbf{B}^* \mathbf{B})^{-1} \mathbf{B}^*$. This function is nonlinear in the parameter $\boldsymbol{\mu}$. The idea of IQML is to instead minimize a series of quadratic cost functions. The estimate of $\boldsymbol{\mu}$ obtained in the k^{th} step of the IQML algorithm is given by

$$\hat{\boldsymbol{\mu}}_k = \arg \min_{\boldsymbol{\mu}^T \boldsymbol{\mu} = 1} V_{\text{IQML}}^k(\boldsymbol{\mu}),$$

$$V_{\text{IQML}}^k(\boldsymbol{\mu}) = \text{Tr}\{\mathbf{B}(\boldsymbol{\mu}) \hat{\mathbf{W}}_k^{-1} \mathbf{B}^*(\boldsymbol{\mu}) \hat{\mathbf{R}}\},$$

where $\hat{\mathbf{W}}_1 = \mathbf{I}$ and

$$\hat{\mathbf{W}}_k = \mathbf{B}^*(\hat{\boldsymbol{\mu}}_{k-1}) \mathbf{B}(\hat{\boldsymbol{\mu}}_{k-1}), \quad k \geq 2. \quad (6)$$

The IQML algorithm works well in many cases, specifically when the noise power is weak. In addition, the computational load of IQML is small when compared to the exact maximum likelihood estimator or the subspace based approaches [7, 8]. However, as pointed out in [6], the algorithm results in biased estimates even when it is initialized with a consistent estimate. This can be seen in Figure 1 where the limiting IQML cost function ($\hat{\mathbf{R}} = \mathbf{R}$) is plotted

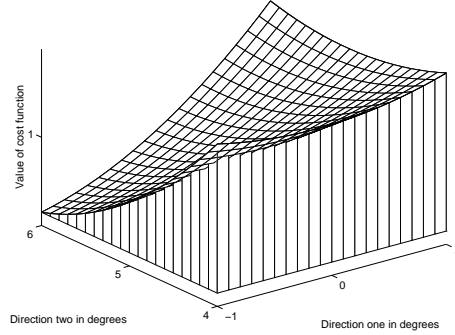


Figure 1: The limiting IQML cost function for the case when two uncorrelated wavefronts impinge from $\boldsymbol{\theta} = [0, 5^\circ]^T$ on a uniform linear array with four sensors. The two sources are equally strong and the signal power of each of them is twice the noise power.

for the case when $\hat{\mathbf{W}}_2$ is calculated using the true parameters. We clearly see that the minima is not at the true directions. Thus, the IQML estimator results in biased estimates.

5. CONSISTENT IQML: MIQML

The MIQML estimator that we propose in this section is identical to IQML in the first step, but modified in the remaining steps so that consistent estimates are obtained. The estimate in the first step is thus given by

$$\hat{\boldsymbol{\mu}}_1 = \arg \min_{\boldsymbol{\mu}^T \boldsymbol{\mu} = 1} V^1(\boldsymbol{\mu}),$$

$$V^1(\boldsymbol{\mu}) = \text{Tr}\{\mathbf{B}(\boldsymbol{\mu}) \mathbf{B}^*(\boldsymbol{\mu}) \hat{\mathbf{R}}\}.$$

Note that, because of the norm constraint and the unitary mapping from $\boldsymbol{\mu}$ to \mathbf{b} , the minimizing argument of this cost function is a consistent estimate of the noise subspace parameters. For a proof of this, see [4]. The norm constraint is crucial for the consistency, because using a linear constraint and the same cost function does *not* yield consistent

estimates. The superior performance of IQML when using a norm constraint has also been observed in [5], but was not exploited to improve the IQML algorithm. In addition, the minimizing value of the cost function is a scaled estimate of the noise power:

$$\hat{\sigma}^2 = \frac{1}{m-d} V^1(\hat{\mu}_1).$$

The above estimate of the noise power is utilized in the subsequent steps of MIQML to modify the IQML criterion as follows

$$\begin{aligned} \hat{\mu}_k &= \arg \min_{\mu^T \mu = 1} V^k(\mu), \\ V^k(\mu) &= \text{Tr}\{\mathbf{B}(\mu) \hat{\mathbf{W}}_k^{-1} \mathbf{B}^*(\mu) (\hat{\mathbf{R}} - \hat{\sigma}^2 \mathbf{I})\}. \end{aligned}$$

The weighting matrix $\hat{\mathbf{W}}_k$ is the same as in the IQML algorithm and is thus given by Equation (6). The subtraction of the noise power guarantees consistency in each step of the algorithm. The minimization of MIQML involves, at least in the first step, the computation of the smallest eigenvalue and the corresponding eigenvector of a $(d+1) \times (d+1)$ matrix. In later steps, the eigenvalue problem can be avoided using a linear constraint on μ . This is due to the subtraction of the noise power in these later steps.

The difference between the IQML and MIQML estimators is thus small, but the performance improvement of the latter is significant in many scenarios. Also observe that the asymptotic performance of MIQML is achieved in the second step. Thus, asymptotically there is no need to perform more iterations. The theoretical performance of the MIQML estimator has been calculated and analytical expressions for the asymptotic covariance of the noise power and the direction estimates can be found in [4]. These investigations show that the MIQML estimates are approximately of the same quality as estimates obtained using DML. However, for closely spaced sources and/or almost coherent sources, there is a performance degradation of the MIQML algorithm compared to the DML estimator. It should be noted that although in those cases the MIQML algorithm does not perform as well as DML, it still performs much better than IQML in all investigated cases (see also Section 7). In Figure 2, the MIQML cost function for the same example as in Section 4 is plotted. Clearly, the minima is in this case located at the true directions.

6. WSF-E

The theoretical investigations of MIQML [4], as well as the simulations, show that the performance of MIQML can differ significantly from the stochastic CRB in cases when the sources are closely spaced and/or they are strongly correlated. Therefore, we now introduce the WSF-E estimator, which is a statistically efficient version of MIQML. The

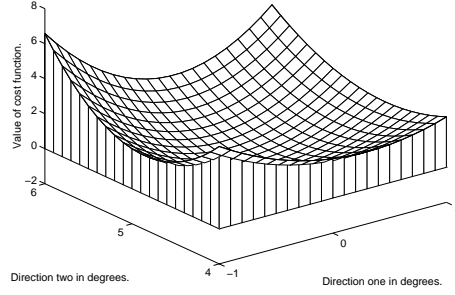


Figure 2: The limiting MIQML cost function corresponding to the example in Section 4.

WSF-E estimator is closely connected to the WSF estimator [8]. The WSF-E cost function is given by

$$\begin{aligned} V_{\text{WSF-E}}(\theta) &= \\ \text{Tr} \left\{ \mathbf{B}(\theta) \hat{\mathbf{W}}_2^{-1} \mathbf{B}^*(\theta) (\hat{\mathbf{R}} - \hat{\sigma}^2 \mathbf{I}) \hat{\mathbf{Q}}_1 (\hat{\mathbf{R}} - \hat{\sigma}^2 \mathbf{I}) \right\}, \end{aligned}$$

where $\hat{\mathbf{W}}_2$ is defined according to (6) and

$$\hat{\mathbf{Q}}_1 = \mathbf{A}(\hat{\theta}_1) \left(\mathbf{A}^*(\hat{\theta}_1) \hat{\mathbf{R}} \mathbf{A}(\hat{\theta}_1) \right)^{-1} \mathbf{A}^*(\hat{\theta}_1).$$

Here, $\hat{\mu}_1$, $\hat{\theta}_1$ and $\hat{\sigma}^2$ are consistent estimates of the corresponding parameters. They can, for example, be obtained from the first step of the MIQML algorithm.

WSF-E inherits several important properties. First, it can be shown that the WSF-E estimates are asymptotically statistically efficient [4]. Second, this optimality is obtained without having to compute an eigendecomposition of the sample covariance matrix, which can be difficult to implement in fixed point arithmetics. Furthermore, by exploiting the structure in the cost function, WSF-E becomes a computationally very attractive algorithm. For example, a square root factorization of $(\hat{\mathbf{R}} - \hat{\sigma}^2 \mathbf{I}) \hat{\mathbf{Q}}_1 (\hat{\mathbf{R}} - \hat{\sigma}^2 \mathbf{I})$ is efficiently obtained by calculating the Cholesky factorization of the $d \times d$ matrix $\mathbf{A}^*(\hat{\theta}_1) \hat{\mathbf{R}} \mathbf{A}(\hat{\theta}_1)$. The estimates of the directions can then be calculated by using a combination of the approaches in [7, 2]. In fact, for large arrays ($m > 8 - 10$), WSF-E requires less computations than IQML. Also, note that MIQML as well as WSF-E still work when the signal covariance matrix is rank deficient [4].

7. SIMULATION RESULTS

Here, we first investigate the proposed MIQML estimator. Data is generated according to (1) and there are two signals present arriving from $\theta^T = [0, 7^\circ]$ relative to the array broadside. The array consists of 8 elements and the signals

are correlated with covariance matrix

$$\mathbf{S} = \begin{bmatrix} 1 & \rho e^{j\pi/4} \\ \rho e^{-j\pi/4} & 1 \end{bmatrix}, \quad (7)$$

where $\rho = 0.5$. The array output covariance matrix is estimated from 1000 snapshots and the results shown in Figure 3 are the average over 100 simulation runs. The number of iterations is three for all the algorithms. The first version

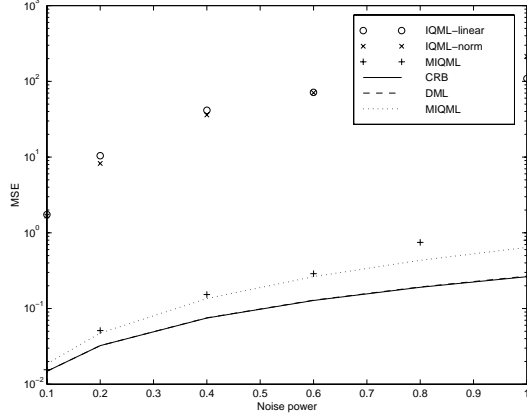


Figure 3: Mean square error for two IQML implementations and MIQML as a function of the noise power. Corresponding theoretical performance is indicated with lines.

of IQML (IQML-linear) is implemented using a linear constraint on the parameter set. The second version of IQML (IQML-norm) is implemented with a norm constraint on the parameters. The proposed algorithm works in agreement with the theoretical expressions. However, note that it performs slightly worse than the theoretical performance of the DML estimator.

We now investigate the WSF-E estimator using almost the same setup. However, now the sources are located at 0 and 5 degrees, respectively, and the signal covariance matrix is in this case given by (7) with $\rho = 0.99$. The results in Figure 4 are obtained when the noise power is equal to 0.8. The WSF-E algorithm is initialized with the estimate from the first step of MIQML, and then the WSF-E algorithm is iterated twice. The optimal WSF-E approach as well as the WSF estimator agrees well with the theoretical results in this case. Note the poor performance of the MIQML estimator for this case and also the difference between the performance of the DML estimator and the stochastic Cramér-Rao lower bound.

8. CONCLUSIONS

With the inconsistency of the IQML algorithm as a motivation, we have developed two consistent IQML like procedures for direction estimation using linear sensor arrays.

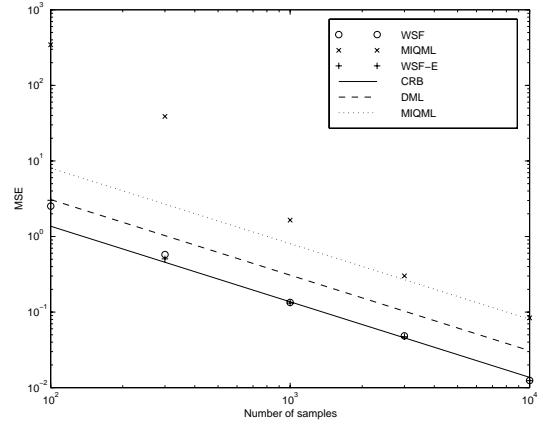


Figure 4: MSE as a function of the number of received snapshots for MIQML, WSF and WSF-E. The corresponding theoretical performance is indicated with lines.

The first algorithm (MIQML) performs close to the deterministic maximum likelihood estimator in many cases. The second algorithm (WSF-E) offers the optimal performance of WSF, but accomplishes this without making use of an eigendecomposition of the sample covariance matrix. The computational complexity of MIQML and WSF-E is close to the complexity of IQML.

9. REFERENCES

- [1] Y. Bresler and A. Macovski. Exact maximum likelihood parameter estimation of superimposed exponential signals in noise. *IEEE Tr. on Acoustics, Speech and Signal Processing*, 34(5):1081–1089, October 1986.
- [2] Y. Hua. The most efficient implementation of the IQML algorithm. *IEEE Tr. on Signal Processing*, 42(8):2203–2204, August 1994.
- [3] H. Krim and M. Viberg. Two decades of array signal processing research: The parameter estimation approach. *IEEE Signal Processing Magazine*, 13(4):67–94, July 1996.
- [4] M. Kristensson, M. Jansson, and B. Ottersten. Modified IQML and weighted subspace fitting without eigendecomposition. Submitted to *Signal Processing*.
- [5] V. Nagesha and S. Kay. On frequency estimation with the IQML algorithm. *IEEE Tr. on Signal Processing*, 42(9):2509–2513, September 1994.
- [6] P. Stoica, J. Li, and T. Söderström. On the inconsistency of IQML. *Signal Processing*, 56(2):185–190, January 1997.
- [7] P. Stoica and K. Sharman. Novel eigenanalysis method for direction estimation. *IEE Proceedings-F*, 137(1):19–26, February 1990.
- [8] M. Viberg and B. Ottersten. Sensor array processing based on subspace fitting. *IEEE Tr. on Signal Processing*, 39(5):1110–1121, May 1991.