PROCEEDINGS OF ISAP '92, SAPPORO, JAPAN

MAXIMUM LIKELIHOOD TECHNIQUE FOR STRUCTURED COVARIANCE ESTIMATION IN ANTENNA ARRAYS

A.B. Gershman, V.I. Turchin, R.A. Ugrinovsky, Institute of Applied Physics Academy of Science, 46 Uljanov Str., 603600, Nizhny Novgorod, Russia

1. Introduction

The problem of the maximum likelihood (ML) estimation of structured covariance has recently been studied in various aspects [1–3]. For antenna array processing supplements, structured ML estimates of complex Toeplitz covariance matrices have been obtained in [4,5]. In [6] the ML estimates of signal and noise powers for the low-rank structured covariance matrix model were derived. In this communication the ML estimates of unknown signal and noise powers for a full-rank structured covariance matrix model are presented. For derivation of our ML estimator we employed the low signal/noise ratio (SNR) assumption. The estimation errors are compared with the Cramer-Rao lower bound.

2. ML estimation of signal and noise powers

Assume, that $n \times n$ noise covariance matrix I and $n \times n$ signal covariance matrix Q are known a priori and, so, the exact covariance matrix of array outputs is given by

$$R = p_n I + p_s Q \ ,$$

where I is the identity matrix, Q is the complex matrix of arbitrary structure, p_s and p_n are unknown signal and noise variances, respectively. The problem is to derive the ML estimates of the real parameters p_s and p_n from the sample covariance matrix $\hat{R}$ given by

$$\hat{R} = \frac{1}{m}\sum_{i=1}^{m} X(i)X^H(i) \ .$$

Here m is the total number of statistically independent array data snapshots, $X(i)$ is jointly Gaussian $n \times 1$ complex array data vector with the following properties: $E\{X(i)\} = 0$, $E\{X(i)X^H(k)\} = \delta_{ik}R$, where δ_{ik} is the Kronecker delta, H denotes the Hermitian transpose, E denotes the expectation operator. The ML function can be expressed as [4]:

$$L = -\log\det R - Tr(R^{-1}\hat{R}) \ , \tag{1}$$

where Tr denotes the trace of matrix. The ML equations can generally be written in the following form:

$$\frac{\partial L}{\partial p_n} = 0; \quad \frac{\partial L}{\partial p_s} = 0 \ , \tag{2}$$

for $p_n = \hat{p}_n$; $p_s = \hat{p}_s$, where $\hat{p}_n$ and $\hat{p}_s$ are the ML estimates of the parameters p_n, p_s. We obtain the solution of (2) under the low SNR assumption given by $p_s\|Q\| \ll p_n$, $p_s \ll p_n$. One can write the expansion:

$$R^{-1} = \frac{1}{p_n}\left(I - \frac{p_s}{p_n}Q + \left(\frac{p_s}{p_n}\right)^2 Q^2 - \dots\right) \ , \tag{3}$$

which converges when the low SNR assumption is fulfilled. Employing (3), we find that

$$\frac{\partial}{\partial p_s}(Tr(R^{-1}\hat{R})) = -\frac{1}{p_n^2}Tr(Q\hat{R}) + 2\frac{p_s}{p_n^3}Tr(Q^2\hat{R}) - \dots \ ,$$

(4)

$$\frac{\partial}{\partial p_n}(Tr(R^{-1}\hat{R})) = -\frac{1}{p_n^2}Tr(\hat{R}) + 2\frac{p_s}{p_n^3}Tr(Q\hat{R}) - 3\frac{p_s^2}{p_n^4}Tr(Q^2\hat{R}) + \dots \ .$$

Rewriting $\det R$ as $\det R = p_s^n \det(\frac{p_n}{p_s}I + Q)$, we find also, that

$$\frac{\partial}{\partial p_s}(\log\det R) = \frac{n}{p_s} - \frac{p_n}{p_s^2}\frac{\sum_{k=1}^{n}\det(Q_k + \frac{p_n}{p_s}I)}{\det(Q + \frac{p_n}{p_s}I)} \ ,$$

(5)

$$\frac{\partial}{\partial p_s}(\log\det R) = \frac{1}{p_s}\frac{\sum_{k=1}^{n}\det(Q_k + \frac{p_n}{p_s}I)}{\det(Q + \frac{p_n}{p_s}I)} \ ,$$

where Q_k is the $(n-1)\times(n-1)$ matrix which is obtained from the matrix Q after destroying its kth row and kth column. After the straightforward calculations we represent (5) in the following form:

$$\frac{\partial}{\partial p_s}(\log\det R) = \frac{1}{p_n}(TrQ - \frac{p_s}{p_n}Tr(Q^2) + \dots) \ ,$$

(6)

$$\frac{\partial}{\partial p_s}(\log\det R) = \frac{1}{p_n}(n - \frac{p_s}{p_n}TrQ + \left(\frac{p_s}{p_n}\right)^2 Tr(Q^2) - \dots) \ ,$$

Note that in general ML equations (2) are nonlinear, but they can be linearized with respect to the small parameter $\frac{p_s}{p_n}$. Using equations (1), (4), and (6) and the above-mentioned linearization procedure we rewrite the ML equations (2) as

$$\frac{1}{p_n}Tr(Q\hat{R}) - \frac{p_s}{p_n}Tr(Q^2) = TrQ \ ,$$

(7)

$$\frac{1}{p_n}Tr(\hat{R}) - \frac{p_s}{p_n}Tr(Q) = n \ .$$

Solution of (7) represents the ML estimates of p_s and p_n:

$$\hat{p}_s \approx \frac{n\ Tr(Q\hat{R}) - TrQ\ Tr\hat{R}}{n\ Tr(Q^2) - (TrQ)^2} \ ,$$

(8)

$$\hat{p}_n \approx \frac{Tr\hat{R}\ Tr(Q^2) - TrQ\ Tr(Q\hat{R})}{n\ Tr(Q^2) - (TrQ)^2} \ .$$

Equations (8) are approximate because of the linearization procedure. The distinction between the ML estimator (8) and the exact solution of equations (2) is derived in the next section.

3. Statistical performance

It is easy to verify that estimates (8) are nonbiased because $E\{\hat{p}_s\} = p_s$, $E\{\hat{p}_n\} = p_n$. Calculation of the covariances of $\hat{p}_s, \hat{p}_n$ yields:

$$cov(\hat{p}_s,\hat{p}_s) = \frac{1}{m}\left(p_n^2 \frac{n}{n\ Tr(Q^2) - (TrQ)^2} + 2p_n p_s \frac{n^2\ Tr(Q^3) - 2n\ TrQ\ Tr(Q^2) + (TrQ)^3}{(n\ Tr(Q^2) - (TrQ)^2)^2} + \right.$$

$$\left. + p_s^2 \frac{n\ Tr(Q^4) - 2n\ TrQ\ Tr(Q^3) + (TrQ)^2\ Tr(Q^2)}{(n\ Tr(Q^2) - (TrQ)^2)^2}\right) ,$$

$$cov(\hat{p}_n,\hat{p}_n) = \frac{1}{m}\left(p_n^2 \frac{Tr(Q^2)}{n\ Tr(Q^2) - (TrQ)^2} + 2p_n p_s \frac{(TrQ)^2\ Tr(Q^3) - TrQ\ (Tr(Q^2))^2}{(n\ Tr(Q^2) - (TrQ)^2)^2} + \right.$$

$$\left. + p_s^2 \frac{(Tr(Q^2))^3 - 2\ TrQ\ Tr(Q^2)\ Tr(Q^3) + (TrQ)^2\ Tr(Q^4)}{(n\ Tr(Q^2) - (TrQ)^2)^2}\right) , \qquad (9)$$

$$cov(\hat{p}_s,\hat{p}_n) = \frac{1}{m}\left(- p_n^2 \frac{TrQ}{n\ Tr(Q^2) - (TrQ)^2} + 2p_n p_s n \frac{(Tr(Q^2))^2 - TrQ\ Tr(Q^3)}{(n\ Tr(Q^2) - (TrQ)^2)^2} + \right.$$

$$\left. + p_s^2 \frac{n(Tr(Q^2)\ Tr(Q^3) - TrQ\ Tr(Q^4)) + (TrQ)^2\ Tr(Q^3) - TrQ(Tr(Q^2))^2}{(n\ Tr(Q^2) - (TrQ)^2)^2}\right) .$$

The estimation losses of (8) as compared to the exact ML solution can be described by means of the following parameter:

$$\eta = \det J \det C \ , \qquad (10)$$

where J is the 2×2 Fisher information matrix and C is the 2×2 covariance matrix of estimates $\hat{p}_s$ and $\hat{p}_n$. This parameter characterizes the distinction between the estimation errors of the proposed estimator and the Cramer-Rao lower bound. It is straightforward to show with respect of (9) that (10) can be expressed as

$$\eta = 1 + 4\left(\frac{p_s}{p_n}\right)^2 \left[\frac{n\ Tr(Q^4) - (Tr(Q^2))^2}{n\ Tr(Q^2) - (TrQ)^2} - \frac{(n\ Tr(Q^3) - TrQ\ Tr(Q^2))^2}{(n\ Tr(Q^2) - (TrQ)^2)^2}\right] + O\left[\left(\frac{p_s}{p_n}\right)^3\right] . \qquad (11)$$

From equation (11) it follows that for low SNR the parameter η is close to unity and, so, the covariances of ML estimates (8) are comparable to the Cramer-Rao lower bound. The second term in (11) gives the SNR threshold for the proposed ML estimator.

4. Example

Consider a simple example where the signal covariance matrix has a diade structure, which corresponds to the case of a localized narrowband signal source. The matrix Q in this case can be expressed as $Q = SS^H$, where S is the $n \times 1$ signal direction vector. Assuming, that the norm of the matrix Q is $S^H S = n$ we obtain from equations (8) that the ML estimates for this special case can be represented as

$$\hat{p}_s = \frac{S^H \hat{R} S - Tr\hat{R}}{n(n-1)} \ ,$$
$$\hat{p}_n = \frac{n\ Tr\hat{R} - S^H \hat{R} S}{n(n-1)} \ . \tag{12}$$

Equations (12) coincide with the exact ML solution since $\eta = 1$. So, the covariances of the ML estimator coincide with the Cramer-Rao lower bound. Using (12) or (9), we find that these covariances are given by

$$cov(\hat{p}_s, \hat{p}_s) = \frac{1}{m}\left(\frac{p_n^2}{n(n-1)} + 2\frac{p_n p_s}{n} + p_s^2\right) \ ,$$
$$cov(\hat{p}_n, \hat{p}_n) = \frac{1}{m}\frac{p_n^2}{n-1} \ ,$$
$$cov(\hat{p}_s, \hat{p}_n) = -\frac{1}{m}\frac{p_n^2}{n(n-1)} \ .$$

From the equations (11) it is seen that in the case of a localized narrowband signal source the signal and noise powers can be optimally estimated using only two simple operations: the steering operation and the measurement of the input power of the array.

5. Conclusion

In this paper we present the novel ML estimator of the signal and noise powers for structured covariance. The proposed estimator is based on the low SNR assumption where the ML equations can be linearized with respect to this small parameter. It is shown that for low SNR the estimation errors are comparable to the Cramer-Rao lower bound.

References

1. R. D. Bock," Components of variance analysis as a structural and discriminal analysis of psychological tests," *Brit. J. Statist. Psych.*, vol.13, pp. 151–163, 1960.

2. T. W. Anderson, "Statistical inference for covariance matrices with linear structure," *Multivariate Analysis II*, P. R. Krishnaian, Ed. New York: Academic, 1969, pp.55-66.

3. P. J. Tourtier and L. L. Scharf, "Maximum likelihood identification of correlation matrices for estimation of power spectra at arbitrary resolutions," in *Proc. ICASSP 87*, Apr. 1987, pp.2066–2069.

4. J. P. Burg, D. G. Luenberger, and D.L.Wenger, "Estimation of structured covariance matrices," *Proc. IEEE*, vol. 70, pp. 963–974, Sept. 1982.

5. H. Wang and G. H. Wakefield, "Signal-subspace approximation for line spectrum estimation," in *Proc. ICASSP 87*, Apr. 1987, pp.2054–2057.

6. B. D. Van Veen and L. L. Scharf, "Estimation of structured covariance matrices and multiple window spectrum analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-38, pp.1467–1472, August 1990.