Optimal Toeplitz Completion of Covariance Matrix for Robust DOA Estimation

Carl-Inge C. Nilsen, *Member, IEEE,* Ines Hafizovic, *Student Member, IEEE*

***Abstract*—Algorithms for direction-of-arrival (DOA) estimation and beamforming will suffer from decreased performance when the estimated statistics deviate from the underlying model. This can be due to estimation from a finite number of data vectors or in case of correlation between signals. Covariance matrix estimators can attempt to counteract these effects by forcing the estimate to have a Toeplitz structure, such as redundancy averaging. It is known that redundancy averaging improves performance, but results in biased DOA estimates that are worse than what can be acomplished with non-Toeplitz constrained techniques such as e.g. spatial averaging. In this letter/correspondence we introduce an optimal, iterative Toeplitz-constrained covariance matrix estimator. We show that the estimator yields redundancy averaging as its first step, and that subsequent steps improve the DOA estimates by reducing the bias and increasing resolution beyond that of redundancy averaging.**

I. Introduction

Several adaptive beamforming and DOA estimation methods are based on examining the spatial covariance matrix. For certain array geometries, such as the uniformly spaced linear array (ULA), the covariance matrix should ideally be Toeplitz. Due to different deviations from our assumptions, such as too few data vectors and/or signal correlation (*should we mention element position errors?*) this criterion is often not satisfied. Several authors have suggested so-called Toeplitz-completion-based estimation methods that force the estimated covariance matrix to be nearly Toeplitz, e.g. adaptive spatial averaging [1], or perfectly Toeplitz, e.g. redundancy averaging [2]. We have chosen to investigate the latter method. It is known that redundancy averaging creates a covariance matrix that is inconsistent with the underlying signal model, yielding biased DOA estimates [3] that can be outperformed by ordinary spatial averaging. It is not obvious how to further develop or modify redundancy averaging to improve its performance, as the technique does not include any parameters and has not been derived using any optimality criterion.

In this correspondence we derive a novel technique for covariance matrix decorrelation through Toeplitz completion and show that a slightly modified form of redundancy averaging is the result of the first iteration. Although the covariance matrix estimates are still inconsistent, the bias induced in the DOA estimates by this inconsistency is significantly reduced for each iteration. The new technique is shown through simulations to have better treshold and resolution properties than both redundancy averaging avd spatial averaging, as well as a DOA RMSE that is superior to redundancy averaging and approaches spatial averaging as SNR increases.

C.-I.C.Nilsen is with the Department of Informatics, University of Oslo. Ines Hafizovic is with Squarehead Technology AS.

Manuscript received XX XX, 2012.

Notes:

- "Decorrelation" (can be other motivations) or "Toeplitz-ization" (underlying causes not immediately clear)? Toeplitz completion....
- Check both ESPRIT and MUSIC?

II. Background

A. Array Model, our Objective [the Spatial Covariance Matrix, and its Estimation]

We assume an M-element uniformly spaced linear array (ULA) which captures N temporal samples of a spatial data vector, $\vec{x}[n]$, consisting of $D \leq M$ plane-wave signals of interest, $\vec{s}[n]$, and spatially white noise, $\vec{n}[n]$:

$$\vec{x}[n] = \vec{s}[n] + \vec{n}[n] \in \mathbb{C}^M, n = 0, 1, \cdots, N-1 \tag{1}$$

The signals of interest can be described as:

$$\vec{s}[n] = e^{i\omega n} \sum_{d=0}^{D-1} A_d \vec{v}_d \tag{2}$$

where A_d are the complex signal amplitudes, and the propagation vectors $\vec{v}_d$ for signals arriving at the ULA from angles θ_k are:

$$\vec{v}_d = \left[e^{-i\frac{2\pi}{\lambda}\sin\theta_d (M-1)/2} \cdots 1 \cdots e^{i\frac{2\pi}{\lambda}\sin\theta_d (M-1)/2} \right]^T . \tag{3}$$

For notational simplicity, we will define the signal wavenumbers as:

$$k_d = \frac{2\pi}{\lambda}\sin\theta_d, \tag{4}$$

and we will also suppress the temporal index n. Our primary goal is to estimate the directions of arrival (DOA) θ_d from the data $\vec{x}[n]$. The number of signals, D, is either known or it must also be estimated from the data.

We define the data spatial covariance matrix of the data as:

$$\mathbf{R} = E\left\{\vec{x}\vec{x}^H\right\} \tag{5}$$

When the signals of interest are uncorrelated, i.e. $E\left\{A_d A_{d+d'}^*\right\} = \sigma_d^2 \delta_{d'}$ or $\vec{v}_d^H \vec{v}_{d+d'} \propto \delta_{d+d'}$, the covariance matrix can be decomposed into the signal and noise covariance matrices:

$$\mathbf{R} = \mathbf{R}_s + \mathbf{R}_n \tag{6}$$
$$\mathbf{R}_s = \sum_{d=0}^{D-1} \sigma_d^2 \vec{v}_d \vec{v}_d^H$$
$$\mathbf{R}_n = \sigma_n^2 \mathbf{I}.$$

These covariance matrices have a Toeplitz structure. When $\mathbf{R}$ is not known, it must be estimated from the data. A "common" (*more sciency definition*) estimator is the sample covariance matrix:

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{n=0}^{N-1} \vec{x}[n] \vec{x}^H[n]. \tag{7}$$

B. DOA Estimation with Root-MUSIC

There are several non-parametric and parametric techniques for achieving our primary goal. The non-parametric methods are typically beamscan algorithms (see e.g. [4]), which form a spatial spectrum over all wavenumbers and locate the D largest peaks within. The most famous parametric methods are MUSIC [5] and ESPRIT [6], and we will concern ourselves with the former (?). Both these methods use the spatial covariance matrix to estimate the DOAs of the D signals. This number D is assumed to be accurately known, which is why the methods are called parametric.

MUSIC uses the signal-plus-noise- and noise-subspace decomposition of the covariance matrix:

$$\mathbf{R} = \mathbf{V} \Lambda \mathbf{V}^H, \mathbf{V} = \left[\mathbf{V}_{S+N}, \mathbf{V}_N\right]. \tag{8}$$

The wavenumbers are the roots of the MUSIC power spectrum:

$$Q_{MUSIC}(k) = \vec{v}^H(k) \mathbf{V}_N \mathbf{V}_N^H \vec{v}(k) \tag{9}$$

C. Robust Covariance Matrix Estimation

In practice, the sample covariance matrix of (7) will not have a Toeplitz structure. This can be due to different reasons. If $\hat{\mathbf{R}}$ is based on a finite number of temporal samples, the deviation from Toeplitz structure will typically increase as the number of samples N is decreased. If the signals of interest are correlated, the sample covariance matrix will always have a non-Toeplitz structure independent of N. An interesting problem is whether it is still possible to regain the Toeplitz covariance matrix of (6) despite these deviations. Several techniques have been suggested, the most well-known being spatial averaging [7]. Spatial averaging consists of dividing the array data vectors $\vec{x}[n]$ into K subarray data vectors $\vec{x}_k[n]$ of length $L = M - K + 1$ that overlap by all but one element:

$$\vec{x}_k[n] = \left[x_k[n], \cdots, x_{k+L-1}[n]\right]^T \in \mathbb{C}^L. \tag{10}$$

The spatially averaged covariance matrix is formed from the subarray data vectors as:

$$\hat{\mathbf{R}}_{SA} = \frac{1}{KN} \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} \vec{x}_k[n] \vec{x}_k^H[n]. \tag{11}$$

Unfortunately, this matrix only achieves Toeplitz structure asymptotically as $K \leftarrow \infty$, which is an unrealistic assumption. Therefore, the alternative method of redundancy averaging was suggested to achieve Toeplitz structure for any value of K and N [2]. Redundancy averaging consists of replacing all elements of the covariance matrix by the average across the corresponding matrix diagonal:

$$[\hat{\mathbf{R}}_{ra}]_{m,n} = \frac{1}{M-P} \sum_{p=0}^{P-1} \left[\hat{\mathbf{R}}\right]_{p,P+p}, P = |m-n| \tag{12}$$

Both spatial and redundancy averaging achieves some degree of robustness against above-mentioned model deviations, and the both have their advantages and disadvantages.

The main advantage of spatial averaging is that it preserves the underlying covariance matrix structure. Therefore, it converges to the true covariance matrix, i.e. $\lim_{K\to\infty} \hat{\mathbf{R}}_{SA} = \mathbf{R}$. The disadvantage is that increasing K decreases the degrees of freedom by the same amount.

The advantage of redundancy averaging is that it can yield a full-rank matrix from even a single data vector, and it does not reduce the degrees of freedom at all. The disadvantage is that it does not preserve the underlying covariance matrix structure, and can therefore not be used to perfectly regain $\mathbf{R}$ except for certain exceptional scenarios [3].

SOMETHING MORE HERE, MOTIVATION, DISCUSSION, ETC.

III. Covariance Matrix Decorrelation Through Optimal Toeplitz Completion

An element of the signal covariance matrix from (6) can be written in the wavenumber domain:

$$\left[\mathbf{R}_s\right]_{m,n} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |S(k)|^2 e^{i(m-n)k} \, dk \tag{13}$$

The wavefield magnitude at wavenumber k_s, $|S(k_s)|^2$, is generally unknown, but can be estimated using beamforming:

$$\left|\hat{S}(k_s)\right|^2 = E\left\{\left|\vec{w}^H \vec{x}\right|^2\right\} = \vec{w}^H \mathbf{R} \vec{w}. \tag{14}$$

We note that the beamformer estimate can be written in the wavenumber domain using a Fourier transformation:

$$\vec{w}^H \mathbf{R} \vec{w} = E\left\{\left|\frac{1}{2\pi} \int_{-\pi}^{\pi} W(k) S(k) \, dk\right|^2\right\}. \tag{15}$$

An optimum estimator of $|S(k)|^2$ is the Minimum Variance Distortionless Response (MVDR) beamformer, which is defined as:

$$\left|\hat{S}_{MV}(k_s)\right|^2 = \vec{w}_{MV}^H \mathbf{R} \vec{w}_{MV} \tag{16}$$
$$\text{where } \vec{w}_{MV} = \text{argmin}_{\vec{w}} \text{VAR}\left\{\vec{w}^H \vec{x}\right\} \text{ s.t. } \vec{w}^H \vec{v}(k_s) = 1.$$

Under the assumption of an uncorrelated wavefield, (16) can be written in the wavenumber domain as in (15):

$$\vec{w}_{MV} = \text{argmin}_{\vec{w}} \frac{1}{2\pi} \int_{-\pi}^{\pi} |W(k)|^2 E\left\{|S(k)|^2\right\} dk \tag{17}$$
$$\text{subject to } W(k_s) = 1 \text{ and } w_m = 0 \text{ for } m < 0 \text{ or } m \geq M$$

Since this expression depends on $|S(k)|^2$, the very quantity we are trying to estimate, we can define an iterative estimator of the uncorrelated wavefield (for steps $e = 1, 2, \cdots, \infty$):

$$\left|\hat{S}^{(e+1)}(k_s)\right|^2 = \vec{w}_{MV}^{(e+1)H}\hat{\mathbf{R}}\vec{w}_{MV}^{(e+1)} \tag{18}$$
$$\text{where } \vec{w}_{MV}^{(e+1)} = \text{argmin}_{\vec{w}} \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|W(k)\hat{S}^{(e)}(k)\right|^2 dk.$$

It is well known that the solution of this MVDR optimization problem is:

$$\vec{w}_{MV}^{(e+1)} = \frac{\hat{\mathbf{R}}^{-(e)}\vec{v}(k)}{\vec{v}^H(k)\hat{\mathbf{R}}^{-(e)}\vec{v}(k)}, \tag{19}$$

where $\hat{\mathbf{R}}^{-(e)}$ is shorthand for $\left(\hat{\mathbf{R}}^{(e)}\right)^{-1}$.This estimate can be used with (13) to specify an iterative Toeplitz covariance matrix estimator of the associated uncorrelated wavefield:

$$\left[\hat{\mathbf{R}}^{(e+1)}\right]_{m,n} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{\vec{v}^H(k)\hat{\mathbf{R}}^{-(e)}\hat{\mathbf{R}}\hat{\mathbf{R}}^{-(e)}\vec{v}(k)}{\left|\vec{v}^H(k)\hat{\mathbf{R}}^{-(e)}\vec{v}(k)\right|^2}e^{i(m-n)k}\,dk. \tag{20}$$

As for any other iterative algorithm, there is the question of intialization. A plausible initial estimate $\left|\hat{S}^{(0)}(k)\right|^2$ requring no knowledge about the situation at hand could be the positive, constant spectrum:

$$\left|\hat{S}^{(0)}(k)\right|^2 = \alpha > 0. \tag{21}$$

However, this automatically yields a closed-form solution for the next step, namely the conventional beamformer, no matter the value of α. Therefore, it makes more sense to initialize with the conventional beamformer directly:

$$\left|\hat{S}^{(0)}(k)\right|^2 = \frac{1}{M^2}\vec{v}^H(k)\hat{\mathbf{R}}\vec{v}(k) \tag{22}$$

An interesting observation comes when inserting (22) into (18) and, subsequently, the result into (20) to estimate $\hat{\mathbf{R}}^{(1)}$, namely that the first covariance matrix estimate becomes (see Appendix A for a proof):

$$[\hat{\mathbf{R}}^{(1)}]_{m,n} = \frac{1}{M^2}\sum_{p=0}^{P-1}\left[\hat{\mathbf{R}}\right]_{p,P+p}, P = |m-n|\,. \tag{23}$$

This is equal to the redundancy averaged estimator of (12) except for a term $-P$ in the denominator of the leading fraction. As we will show in the next section, both this scaling and iterating further will improve the performance beyond that of redundancy averaging. However, since step 1 of the iteration still yields a closed form solution, it ultimately makes more sense to start with (23) instead of the conventional beamformer estimate.

Unfortunately, it is not trivial to prove the convergence of the iterative estimator in (20). In the next section, we will demonstrate convergence for a range of cases through simulations.

It is worth noting that the suggested method is more computationally demanding than both spatial and redundancy averaging. Each iteration requires the formation of an $M \times M$ Toeplitz matrix, where the value along each of the M diagonals is calculated from the integral in (20). As there is no closed form solution to this integral, it must be done numerically. The resulting covariance matrix must be then be inverted, which has complexity of $O(M^2)$ or less due to its Toeplitz structure.

IV. Simulations and Discussion

In this section we will compare the performance of redundancy averaging, spatial averaging, and the suggested method for DOA estimation using the Root-MUSIC algorithm.

V. Conclusion

We have introduced an iterative, Toeplitz-constrained covariance matrix estimator for use in parametric DOA estimation. It is shown that this estimator has a slightly modified redundancy average estimator as its first solution, while the subsequent solutions result in improved DOA estimates and resolution.

Appendix A: Proof of Equation (23)

Because the covariance matrix of (13) is Toeplitz, we will investigate the value along the l^{th} diagonal:

$$\begin{aligned}
\left[\hat{\mathbf{R}}^{(1)}\right]_{m,n} &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\hat{S}^{(0)}(k)\right|^2 e^{ilk}\,dk, \\
&\text{for } l \triangleq m-n \\
&= \frac{1}{2\pi M^2}\int_{-\pi}^{\pi}\vec{v}^H(k)\hat{\mathbf{R}}\vec{v}(k)e^{ilk}\,dk \\
&= \frac{1}{2\pi T M^2}\sum_{n=0}^{T-1}\sum_{m'=0}^{M-1}\sum_{n'=0}^{M-1}x_{m'}[t]x_{n'}^*[t]\int_{-\pi}^{\pi}e^{i(l-l')k}\,dk, \\
&\text{for } l' \triangleq m'-n' \\
&= \frac{1}{M^2}\sum_{m'=0}^{M-1}\sum_{n'=0}^{M-1}\delta_{l-l'}\frac{1}{T}\sum_{t=0}^{T-1}x_{m'}[t]x_{n'}^*[t] \\
&= \frac{1}{M^2}\sum_{p=0}^{l-1}\left[\hat{\mathbf{R}}\right]_{p,p+l},
\end{aligned} \tag{24}$$

which we confirm is equal to (23) and nearly equal to the redundancy averaged estimate of the covariance matrix as given in (12), except for the denominator in the leading fraction.

References

[1] K. Takao and N. Kikuma, "An adaptive array utilizing an adaptive spatial averaging technique for multipath environments," *Antennas and Propagation, IEEE Transactions on*, vol. 35, no. 12, pp. 1389 – 1396, dec 1987.

[2] L. Godara, "Beamforming in the presence of correlated arrivals using structured correlation matrix," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 38, no. 1, pp. 1 –15, jan 1990.

[3] K. Indukumar and V. Reddy, "A note on redundancy averaging," *Signal Processing, IEEE Transactions on*, vol. 40, no. 2, pp. 466 –469, feb 1992.

[4] H. V. Trees, *Optimum Array Processing*. Wiley-Interscience, 2002.

[5] R.O.Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation, Vol. AP-34, pp. 276–280*, 1986.

[6] R. Roy, "Esprit: Estimation of signal parameters via rotational invariance techniques)," *Ph.D. Dissertation, Stanford University, Stanford,California*, 1987.

[7] T.-J. Shan and T. Kailath, “Adaptive beamforming for coherent signals and interference,” *IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-33, pp. 527–534*, 1985.