The Application of Improved Covariance Estimation to Adaptive Beamforming and Detection*

Douglas A. Abraham

Naval Undersea Warfare Center
New London, CT 06320
(203) 440-5057
Fax: (203) 440-4010

Dipak K. Dey

Department of Statistics
University of Connecticut
Storrs, CT 06269-3120
(203) 486-4192

Abstract

Adaptive beamformers often rely on estimates of the interference covariance structure. When minimal data is used in estimating the unknown covariance matrix, performance may be improved by modifying the eigenvalues of the sample covariance matrix. The orthogonally invariant covariance estimators proposed by Dey and Srinivasan [1] and a constant risk, minimax estimator are applied to adaptive beamforming by considering the signal-to-interference ratio (SIR) loss factor of Reed, Mallett, and Brennan [2], where the average SIR loss is seen to decrease, and, to adaptive detection by considering the adaptive detectors of Kelly [3] and Robey et. al. [4] where detection performance is seen to improve.

1 Introduction

Adaptive beamforming algorithms rely on estimates of the interference covariance structure to provide improved detection and estimation performance over conventional beamforming. The data used to estimate the unknown interference covariance structure is commonly called auxiliary data. When minimal auxiliary data is used, the variance in the estimates severely degrades performance—potentially providing worse performance than conventional beamforming. Reed, Mallet, and Brennan [2] indicate that, for complex Gaussian data vectors, $2n$ samples of auxiliary data in a maximum likelihood estimate of the n-by-n interference covariance matrix $\boldsymbol{\Sigma}$ are required to achieve less than a 3 dB average loss in the beam output signal-to-interference ratio (SIR). Kelly [3] and Robey et. al. [4] have illustrated the improvement attainable in array signal detection performance by increasing the amount of auxiliary data. For large n, adequate performance could require so much auxiliary data that the interference may become non-stationary. As illustrated by Gray et. al. [5] and Fuhrmann [6] for toeplitz covariance matrices, adaptive array performance improvement may also be obtained by assuming that $\boldsymbol{\Sigma}$ has a specific structure and exploiting this structure to provide improved covariance estimation.

When there is not enough data to adequately estimate $\boldsymbol{\Sigma}$, or, when no specific covariance structure may be assumed, the interference covariance matrix estimate may still be improved by employing a minimax estimator or by applying shrinkage-expansion (S-E) techniques to the eigenvalues of the sample covariance matrix. It is known that the eigenvalues of the sample covariance matrix, a scale of a complex Wishart distributed matrix, are biased away from the true eigenvalues. To a first order approximation, the mean of the larger (smaller) sample eigenvalues is greater (less) than the corresponding true eigenvalues. As shown by Dey and Srinivasan [1], applying S-E techniques that shift the sample eigenvalues toward their geometric mean can result in covariance estimators that dominate the sample covariance matrix in terms of risk under Stein's loss function. The form of the orthogonally invariant, scale invariant, minimax covariance estimators of [1], and a discussion of how to choose two constants associated with the S-E term, is found in section 2.

The utility of these improved covariance estimators in adaptive beamforming and detection is explored in sections 3 and 4, respectively, by considering Reed, Mallet, and Brennan's SIR loss factor and the adaptive detectors of Kelly and Robey.

*This work was sponsored by the Office of Naval Research. Approved for public release; distribution is unlimited.

1058-6393/95 $4.00 © 1995 IEEE

2 Improved Covariance Estimation

In many situations in adaptive beamforming and detection, it is reasonable to assume that the observed data follow a zero mean, multivariate, complex Gaussian distribution,

$$\mathbf{x}_i \sim \mathcal{CN}_p(\mathbf{0}, \mathbf{\Sigma}), \tag{1}$$

for $i = 1, \ldots, k$, where the notation $\mathbf{x} \sim \mathcal{CN}_p(\mu, \mathbf{\Sigma})$ indicates that the p-by-1 vector $\mathbf{x}$ is complex Gaussian distributed with mean μ and covariance $\mathbf{\Sigma}$. The maximum likelihood estimate of $\mathbf{\Sigma}$ is the sample covariance matrix,

$$\hat{\mathbf{\Sigma}}_0 = \frac{1}{k}\sum_{i=1}^{k} \mathbf{x}_i \mathbf{x}_i^H, \tag{2}$$

where the superscript ^{H} represents the complex conjugate and transpose operation. Let the eigenvalue-eigenvector decomposition of the sample covariance matrix be

$$\hat{\mathbf{\Sigma}}_0 = \mathbf{MLM}^H, \tag{3}$$

where $\mathbf{M}$ is the orthogonal matrix of the sample eigenvectors and $\mathbf{L}$ is a diagonal matrix of the sample eigenvalues $(l_1 > l_2 > \cdots > l_p)$.

The orthogonally invariant, scale invariant covariance estimators of Dey and Srinivasan [1] may be described by the form

$$\hat{\mathbf{\Sigma}} = \mathbf{M}\phi(\mathbf{L})\mathbf{M}^H, \tag{4}$$

where the function ϕ operates on the diagonal elements of $\mathbf{L}$ according to

$$\phi_i(\mathbf{L}) = a_i l_i - \frac{c l_i}{b+u}\log\left(\frac{l_i}{\bar{l}}\right), \tag{5}$$

for $i = 1, \ldots, p$. Here, $\bar{l} = \sqrt[p]{\prod_{i=1}^{p} l_i}$ is the geometric mean of the sample eigenvalues and

$$u = \sum_{i=1}^{p}\left[\log\left(\frac{l_i}{\bar{l}}\right)\right]^2. \tag{6}$$

Four covariance estimators of this form will be considered: (i) the maximum likelihood estimator (MLE) of $\mathbf{\Sigma}$ ($a_i = 1$ and $c = 0$), (ii) the MLE with S-E ($a_i = 1$ and $c > 0$), (iii) a minimax estimator, say $\hat{\mathbf{\Sigma}}_m$, with constant risk with respect to $\mathbf{\Sigma}$ ($a_i = \frac{k}{k+p+1-2i}$ and $c = 0$), and (iv) a minimax estimator with S-E ($a_i = \frac{k}{k+p+1-2i}$ and $c > 0$).

Dey and Srinivasan showed that the risk under Stein's loss function[1],

$$L\left(\hat{\mathbf{\Sigma}}, \mathbf{\Sigma}\right) = \mathrm{tr}\left(\hat{\mathbf{\Sigma}}\mathbf{\Sigma}^{-1}\right) - \log\left|\hat{\mathbf{\Sigma}}\mathbf{\Sigma}^{-1}\right| - p, \tag{7}$$

is reduced for the S-E estimators of equation (5) when

$$c < \frac{12(p-3)}{5\tilde{k}} \tag{8}$$

and

$$\sqrt{b} > c, \tag{9}$$

where

$$\tilde{k} = \begin{cases} k & \text{when} \quad a_i = 1 \\ k+p-1 & \text{when} \quad a_i = \frac{k}{k+p+1-2i} \end{cases}. \tag{10}$$

In general, the constant c may be replaced by any non-decreasing function of u that satisfies equation (8). It should be noted that the work of Dey and Srinivasan [1] was strictly for estimating the covariances of real multivariate Gaussian random vectors. Their proofs were based on algebraic arguments applied to the eigenvalues. Thus, their results may be directly extended to the complex Gaussian case where the eigenvalues are also real due to the complex conjugate symmetry of covariance matrices.

The analysis of Dey and Srinivasan [1] does not preclude a risk reduction for values of b and c not satisfying (8) and (9), and, unfortunately, does not provide a method for choosing b and c to minimize the risk. The risk obtained by choosing $b = \gamma_b b_{min}$ and $c = \gamma_c c_{max}$, where $c_{max} = \frac{12(p-3)}{5\tilde{k}}$ and $b_{min} = c_{max}^2$, is shown as a function of γ_b and γ_c, respectively, in figures 1 and 2 for $p = 8$, $k = 16$, and $\mathbf{\Sigma} = \mathbf{I}_p$, the p dimensional identity matrix. In figure 1, $c = c_{max}$, and, in figure 2, $b = b_{min}$. Observe that, as expected, the risk for the S-E methods tends toward that for the non-S-E when either $b \to \infty$ or $c \to 0$. The values of γ_b and γ_c that minimize the risk are functions of $\mathbf{\Sigma}$, and thus may not be chosen without prior knowledge. However, over a wide variety of $\mathbf{\Sigma}$ and k, the trends observed indicated that choosing $\gamma_c = 4$ for the MLE S-E estimator and $\gamma_c = 3$ for the minimax S-E estimator resulted in a risk reduction. The amount of reduction diminishes, potentially becoming negative, as the eigenvalue spread increases, a result consistent with that reported by Dey and Srinivasan [7]. As seen in figure 1, scaling b toward zero reduces the risk, however only slightly. Note that

[1]The use of Stein's loss function may be justified through recognizing that the resulting risk function contains terms found in the entropy function for $\hat{\mathbf{\Sigma}}_0$. Thus, reducing the risk may reduce the entropy associated with estimating $\mathbf{\Sigma}$.

the minimax estimator with S-E obtains the smallest risk. Although some robustness has been observed for varying $\mathbf{\Sigma}$, p, and k, blind use of these values for γ_b and γ_c may result in performance deterioration. Further research is required to provide rules for choosing values of γ_b and γ_c that yield adequate performance for diverse covariance structures.

3 Adaptive Beamforming Performance

The primary goal of adaptive beamforming is to maximize the signal-to-interference ratio at the beam output subject to a constraint forcing signals from a specified direction to be distortionless. When the interference covariance matrix ($\mathbf{\Sigma}$) is perfectly known, the maximum beam output SIR is achieved by using the well known minimum variance distortionless response beamforming weight vector. However, when $\mathbf{\Sigma}$ is estimated by $\hat{\mathbf{\Sigma}}$, a loss is incurred. Conditioned on $\hat{\mathbf{\Sigma}}$, the achieved SIR has the form

$$s = \rho s_{max}, \tag{11}$$

where s_{max} is the maximum achievable SIR,

$$\rho = \frac{\left(\mathbf{d}^H \hat{\mathbf{\Sigma}}^{-1} \mathbf{d}\right)^2}{\left(\mathbf{d}^H \hat{\mathbf{\Sigma}}^{-1} \mathbf{\Sigma} \hat{\mathbf{\Sigma}}^{-1} \mathbf{d}\right)\left(\mathbf{d}^H \mathbf{\Sigma}^{-1} \mathbf{d}\right)} \tag{12}$$

is the SIR loss factor, and $\mathbf{d}$ represents an array steering vector. Reed, Mallett, and Brennan [2] found that ρ followed a beta probability distribution when the sample covariance matrix of complex Gaussian data was used ($\hat{\mathbf{\Sigma}} = \hat{\mathbf{\Sigma}}_0$) irrespective of the structure of $\mathbf{\Sigma}$.

Following Fuhrmann [6], the SIR loss factor will be used to determine the potential improvement in adaptive beamforming due to improved estimation of the interference covariance matrix. Histograms of ρ for the four covariance estimators described in section 2 are found in figure 3 for $p = 8$ sensors equi-spaced in a line array at the design frequency, $k = 16$ samples of auxiliary data, a broadside look direction ($\mathbf{d} = \frac{\mathbf{1}_p}{\sqrt{p}}$), and spatially white noise ($\mathbf{\Sigma} = \mathbf{I}_p$). Note that the histogram for ρ with the sample covariance matrix estimator (MLE) closely follows the predicted beta density function and that the other covariance estimators all show more mass at higher values of ρ, thus, indicating improved adaptive beamforming performance. The effect of these estimators on angle of arrival estimation (e.g., bias and variance) has not been determined.

The expected value of ρ is used as a performance measure to evaluate the adaptive beamforming performance as the amount of auxiliary data or the covariance structure varies. In figure 4 it is seen that less than a one dB loss in the average SIR may be obtained for auxiliary data sizes $k \geq 12$ and $k \geq 15$, respectively, for the minimax S-E and MLE S-E covariance estimators and for $k \geq 20$ for the minimax estimator. The MLE estimator requires $k \geq 34$ to achieve the same performance, nearly three times that of the minimax S-E estimator. The amount of improvement due to the S-E techniques diminishes as the amount of auxiliary data increases.

In figure 5, $\mathrm{E}[\rho]$ is shown as a function of the sensor level interference-to-noise power ratio (INR), s_0, where $\mathbf{\Sigma} = s_0 \mathbf{d}_0 \mathbf{d}_0^H + \mathbf{I}_p$ and $\mathbf{d}_0$ is the steering vector for an interfering signal at 5 degrees from broadside to the array. As the interference power increases, the improvement of the S-E estimators diminishes, a result consistent with the results of section 2 where risk improvement decreased as condition number increased. The minimax estimator with S-E actually provided worse performance than that for the non-S-E method at the higher INR, exhibiting the need for more appropriate choices for the b and c constants used in the eigenvalue modification.

4 Adaptive Detection Performance

Both Kelly [3] and Robey et. al. [4] have proposed constant false alarm rate adaptive detectors. Their hypothesis test assumed that auxiliary data was available to estimate $\mathbf{\Sigma}$ and that the data vector $\mathbf{x}$ was to be tested for the presence of a deterministic signal with unknown complex amplitude. This results in assuming that

$$\mathbf{x} \sim \mathcal{CN}_p(\theta \mathbf{d}, \mathbf{\Sigma}), \tag{13}$$

where $\mathbf{d}$ is the array steering vector pointing in the direction of arrival of the signal and θ is the unknown complex signal amplitude. Kelly's generalized likelihood ratio (GLR) test has the detection statistic

$$T_{GLR} = \frac{\left|\mathbf{d}^H \hat{\mathbf{\Sigma}}^{-1} \mathbf{x}\right|^2}{\left(\mathbf{d}^H \hat{\mathbf{\Sigma}}^{-1} \mathbf{d}\right)\left(1 + \mathbf{x}^H \hat{\mathbf{\Sigma}}^{-1} \mathbf{x}\right)}. \tag{14}$$

Robey's adaptive matched filter (AMF) has the detection statistic

$$T_{AMF} = \frac{\left|\mathbf{d}^H \hat{\mathbf{\Sigma}}^{-1} \mathbf{x}\right|^2}{\left(\mathbf{d}^H \hat{\mathbf{\Sigma}}^{-1} \mathbf{d}\right)}. \tag{15}$$

The probability of detection as a function of the SIR ($s = |\theta|^2 \mathbf{d}^H \mathbf{\Sigma}^{-1} \mathbf{d}$) for Kelly's GLRT and Robey's

AMF are found, respectively, in figures 6 and 7 for detectors formed from the covariance estimators of section 2 for a false alarm probability of $P_{FA} = 10^{-2}$, $p = 8$ sensors in the same configuration as section 3, $k = 16$, and $\mathbf{\Sigma} = \mathbf{I}_p$. The performance of a GLR detector that has knowledge of the interference covariance matrix (but not the complex signal amplitude), known as the clairvoyant detector, is also depicted. Observe that the detectors based on the minimax and S-E covariance estimators show improved performance over the one based on the sample covariance matrix. The improvement for Robey's AMF detector was slightly greater than that for Kelly's GLR detector.

It is important to note that the only detectors having a constant false alarm rate are those using just the sample covariance matrix estimator. The effect of the minimax covariance estimator and the S-E estimators on the false alarm performance for varying $\mathbf{\Sigma}$ has not been evaluated.

5 Conclusions

Adaptive beamforming and detection techniques rely on estimates of the unknown interference covariance matrix to provide improved performance over conventional methods. When minimal auxiliary data is used in the estimation, the variability of the covariance estimate can substantially degrade performance. It has been shown that the adaptive beamforming and detection performance loss due to estimation of an unstructured covariance matrix may be reduced by modifying the eigenvalues of the sample covariance matrix. The performance improvement diminishes as the amount of auxiliary data used to estimate the covariance matrix increases, or, as the eigenvalue spread of $\mathbf{\Sigma}$ increases.

Several issues remain unresolved. The bounds on the b and c constants in the eigenvalue S-E techniques that guarantee risk reduction are not tight. Methods for choosing b and c to minimize the risk for a wide range of covariances are needed. The effects of the S-E and minimax estimators on beamforming applications such as signal waveform or direction of arrival estimation need to be explored. The application of the improved eigenvalue estimation techniques to adaptive detection can not be endorsed unless constant false alarm rate detectors are not imperative. Alternatively, if the false alarm probability of the detectors utilizing the eigenvalue modification techniques can be upper bounded over all possible covariances, improved detection performance may then be obtainable with thresholds that do not change with $\mathbf{\Sigma}$.

References

[1] D. K. Dey and C. Srinivasan, "Estimation of a Covariance Matrix Under Stein's Loss," *The Annals of Statistics*, vol. 13, no. 4, pp. 1581–1591, 1985.

[2] I. S. Reed, J. D. Mallet, and L. E. Brennan, "Rapid Convergence Rate in Adaptive Arrays," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-10, no. 6, pp. 853–863, Nov. 1974.

[3] E. J. Kelly, "An Adaptive Detection Algorithm," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-22, no. 1, pp. 115–127, March 1986.

[4] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg, "A CFAR Adaptive Matched Filter Detector," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-28, no. 1, pp. 208–216, January 1992.

[5] D. Gray, B. Anderson, and P. Sim, "Estimation of Structured Covariances with Application to Array Beamforming," *Circuts, Systems and Signal Processing*, vol. 6, no. 4, pp. 421–447, 1987.

[6] D. R. Fuhrmann, "Application of Toeplitz Covariance Estimation to Adaptive Beamforming and Detection," *IEEE Trans. on Signal Processing*, vol. 39, no. 10, pp. 2194–2198, October 1991.

[7] D. K. Dey and C. Srinivasan, "Estimation of Covariance Matrix Under Stein's Loss," Tech. Rpt., Texas Tech. University, 1984.

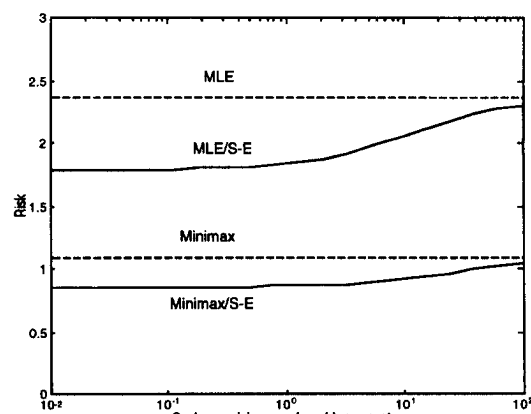

Figure 1: Risk as a function of γ_b with $c = c_{max}$.

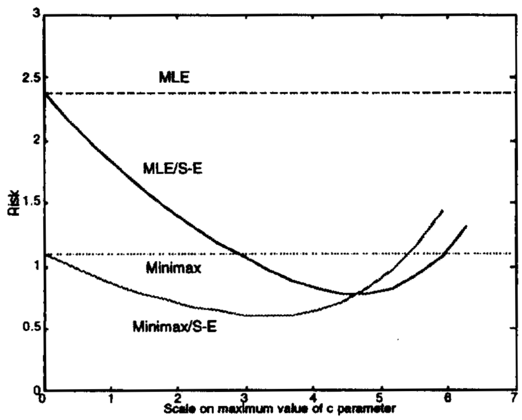

Figure 2: Risk as a function of γ_c with $b = b_{min}$.

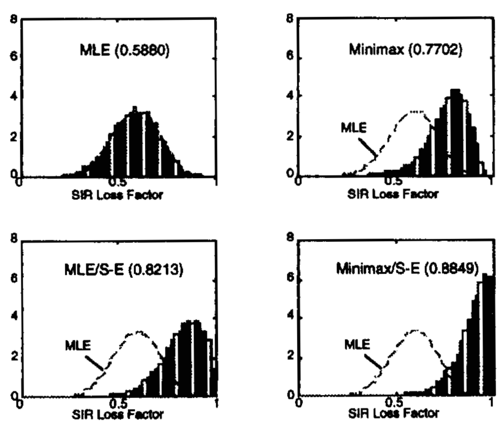

Figure 3: Normalized histogram for SIR loss factor. Numbers in parentheses are the sample mean values. Dashed line is the PDF for ρ when the MLE for $\mathbf{\Sigma}$ is used, with mean value $\mathrm{E}[\rho] = 0.5882$.

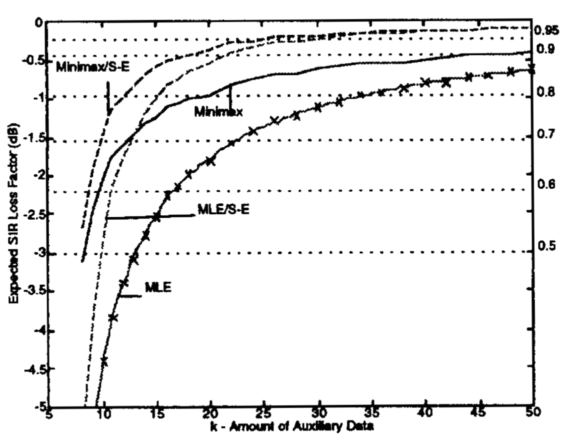

Figure 4: $\mathrm{E}[\rho]$ as a function of the amount of auxiliary data.

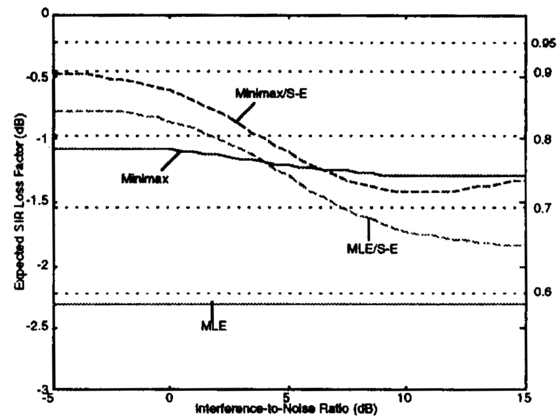

Figure 5: $\mathrm{E}[\rho]$ as a function of interference-to-noise ratio.

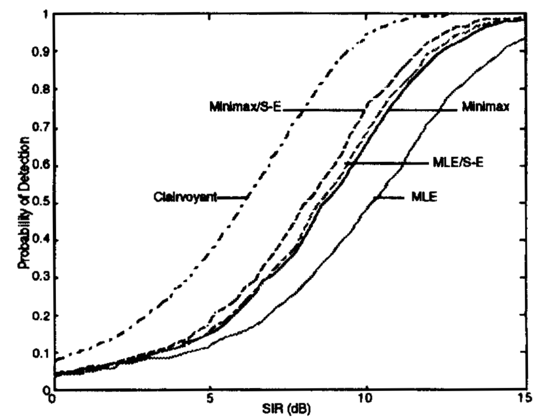

Figure 6: Probability of detection for Kelly's GLRT detector and the clairvoyant detector.

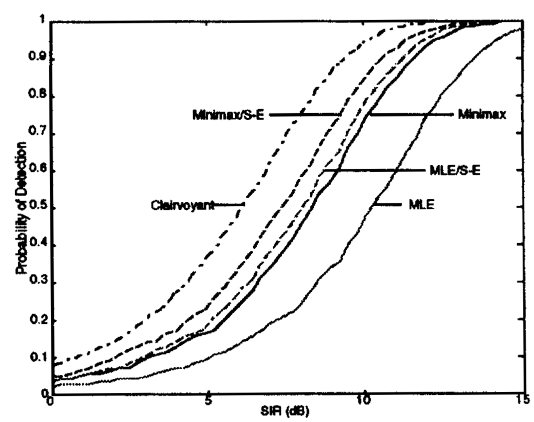

Figure 7: Probability of detection for Robey's AMF detector and the clairvoyant detector.