Robust Krylov-Subspace Methods for Passive Sonar Adaptive Beamforming

By Samuel D. Somasundaram and Nigel H. Parsons
General Sonar Studies, Thales Underwater Systems, Cheshire, UK
(sdsomasundaram@hotmail.com, nigel.parsons@uk.thalesgroup.com)

Abstract

Krylov-subspace methods, such as the multistage Wiener filter and conjugate gradient method, are often used for reduced-dimension adaptive beamforming. These techniques do not, however, allow for steering vector mismatch, which is typically present in many applications of interest, including passive sonar. Here, we discuss recently proposed robust methods that do allow for steering vector mismatch by combining Krylov-subspace methods with robust Capon beamforming techniques. We present experimental data examples that illustrate the benefits of using these algorithms in passive sonar.

1. Introduction

In passive sonar, arrays of hydrophones are employed to sense the sound pressure field for the purpose of detecting and locating sources of sound in the ocean. Beamforming is used to process the hydrophone outputs to form receive beams steered in a chosen direction. Conventional delay-and-sum (DAS) beamforming steers the beam to a specific direction by delaying the hydrophone outputs so that a signal received from a source in the specified direction will appear aligned in the delayed hydrophone outputs. Often, frequency-domain beamformers are implemented, where delays are produced by applying a set of frequency-dependent phase shifts. Summation of the delayed outputs results in coherent summation of the source signal, but not noise that is uncorrelated between hydrophones, leading to an increase in signal-to-noise ratio (SNR). It is well-known that DAS beamforming is only optimal for a single source in uncorrelated noise (see, e.g., [Stoica & Moses 2005]). In passive sonar, there are typically multiple sources of correlated noise, including contacts not in the steer direction, ambient isotropic noise, platform noise, e.g., from the vessel on which the array is mounted, and flow and flow-induced noises. Shading can be used with DAS beamformers to trade mainbeam width for low sidelobe level. Apart from the fact one might not want to make such a trade, faulty sensors and/or differences in sensor calibration often result in the desired specifications (e.g., sidelobe level) not being fully achieved.

Alternatively, adaptive beamforming can be used to produce data-dependent weights that can direct nulls towards sources of noise to increase SNR. Commonly, frequency-domain implementations of the minimum variance distortionless response (MVDR) or Capon beamformer and its variants are used. In MVDR beamforming, weights that minimise the array output power subject to the constraint that they pass a signal array steering vector (ASV) undistorted are sought. Weights are designed for each frequency bin and steer direction. The ASV models the array response to a unit signal at the frequency and steer direction of interest. Minimising the array output power enables re-
projection of noise, whereas the constraint on the ASV protects the signal-of-interest (SOI) in the assumed steer direction. The optimal MVDR weights are a function of the inverse array covariance matrix and the assumed SOI ASV. Unfortunately, several issues plague the implementation of the MVDR/Capon beamformer in practice. The first is the sensitivity to differences in the assumed and actual ASVs, resulting from, e.g., angle-of-arrival or pointing errors, sensor calibration errors, source wavefront distortions (due to inhomogeneities in the ocean) and scattering, all of which leads to SOI cancellation. Secondly, the array covariance matrix is unknown in practice and must be estimated from available data. The MVDR weights are highly sensitive to errors in the covariance matrix estimate. Generally, the larger the covariance matrix, the more data is required to obtain a satisfactory covariance matrix estimate, which leads to problems implementing the MVDR beamformer on large arrays. The third difficulty arises in the need to compute the inverse of the array covariance matrix, which requires a computational complexity cubic in the number of degrees of freedom. Many robust adaptive beamforming methods have been proposed to alleviate these problems and this area continues to be an active area of research.

Robust Capon beamforming and its variants [Vorobyov et al. 2003, Stoica et al. 2003, Lorenz & Boyd 2005, Li & Stoica 2005] exploit ellipsoidal ASV uncertainty sets to systematically deal with ASV mismatch and have been successfully applied in passive sonar [Somasundaram & Parsons 2011, Somasundaram 2011, Somasundaram et al. 2012, Somasundaram 2013]. Instead of protecting only a single SOI ASV, they protect any ASV within the ellipsoidal set. When exploiting spherical ASV uncertainty (which is a special case of ellipsoidal uncertainty), the RCB also exhibits good robustness to covariance matrix errors; however, there is still room for improvement in this area, especially when beamforming large arrays operating in dynamic environments and/or when exploiting non-spherical uncertainty sets. Further, the RCB complexity is still cubic in the number of adaptive degrees of freedom and would benefit from computational efficiencies. One approach is to exploit dimensionality reduction to speed up convergence and reduce complexity. In [Somasundaram 2011], the reduced-dimension robust Capon beamforming (RDRCB) framework for combining RCB and reduced-dimension techniques was proposed, in which a (complex) propagation theorem for deriving reduced-dimension uncertainty sets from full-dimension (element-space) uncertainty sets and dimension reducing transforms was derived. In [Somasundaram 2011], data-independent beamspace dimensionality reduction techniques were exploited in the framework and shown to provide benefits in passive sonar beamforming. Recently, data-dependent Krylov-subspace techniques have been considered within the framework [Somasundaram et al. 2013, Somasundaram et al. (submitted)]. Krylov-subspace techniques appear in many well-known reduced-dimension algorithms. For instance, the multistage Wiener filter (MSWF) [Goldstein et al. 1998] can be implemented by using a Krylov-subspace basis, formed using the (non-orthogonal) Powers-of-R (NO-PoR) method, to project the data onto the Krylov subspace and then finding the optimal reduced-dimension Wiener filter within this subspace. In [Ge et al. 2006, Kirsteins & Ge 2006], an orthogonal Powers-of-R (O-PoR) method was used to form a Krylov-basis, which was used to reduce dimensions before finding the optimal reduced-dimension MVDR beamformer. Further, the conjugate-gradient (CG) method [Hestenes & Stiefel 1952], which is used to solve quadratic minimisation problems such as the MVDR problem, operates in a Krylov-subspace. The aforementioned Krylov-subspace techniques, which all expand the same subspace, do not allow for ASV mismatch and are sensitive to model-order over-determination.

In [Somasundaram et al. 2013, Somasundaram et al. (submitted)], Krylov-subspace dimensionality reduction based on the NO-PoR, O-PoR and CG algorithms was exploited in
the RDRCB framework to produce PoR and CG-based Krylov-subspace RDRCBs, which were shown to provide robustness to ASV mismatch and model over-determination. The computational complexity analysis in [Somassundaram et al. (submitted)] indicated that the CG-based RDRCBs are less complex than the PoR based ones and, for an $M = 320$ element array, the CG-RDRCB under spherical uncertainty and with reduced-dimension $N = 5$ is around 900 times less complex than the (element-space) RCB under spherical uncertainty. Further, CG-RDRCBs exploiting spherical or non-degenerate uncertainty sets are only 1.2 and 2.2 times more complex than the standard non-robust CG-MVDR algorithm. Since the NO-PoR, O-PoR and CG algorithms all expand the same Krylov-subspace, one might expect them to give identical results, however, it was found that the CG and O-PoR versions were numerically more stable than the NO-PoR based method.

In [Somassundaram et al. (submitted)], additional enhancements were proposed by combining the Krylov-subspace RDRCBs with stopping criteria, which prevent the model order from being selected too high. When exploiting the stopping criteria, there was no noticeable differences in the performances of the Krylov-subspace RDRCBs. So far, the Krylov-based RDRCBs have only been evaluated on simulated array data. Here, we present preliminary results for the CG-RDRCB, exploiting a stopping criterion, on recorded passive sonar data, comparing it to the DAS and CG-MVDR beamformers.

2. Data Modelling

The time domain output of the $m$th sensor $X_m(t)$ can be written as $X_m(t) = S_0(t - \tau_m(\theta_0)) + N_m(t)$, where $S_0(t)$, $\tau_m(\theta_0)$ and $N_m(t)$ denote the desired signal waveform, the propagation delay to the $m$th sensor, relative to some reference point, for the desired signal impinging from a location described by $\theta_0$, and the noise-plus-interference at the $m$th sensor, which is assumed to be uncorrelated with the desired signal. An $L$-point FFT is used to transform the sensor time-series data to the frequency-domain. For the $l$th FFT bin, $m$th element and $\ell$th frequency-domain snapshot, we write

$$x_{m,k}(f_l) = s_{0,k}(f_l)e^{-i2\pi f_l\tau_m(\theta_0)} + n_{m,k}(f_l), \quad l = 1 \ldots L$$

with $x_{m,k}(f_l)$, $s_{0,k}(f_l)$ and $n_{m,k}(f_l)$ denoting the Fourier transform, at frequency $f_l = (l-1)/LT_s$ with $T_s$ denoting the sampling period, of $X_m(t)$, $S_0(t)$ and $N_m(t)$, respectively. The array steering vector model for a signal with center frequency $f_1$ impinging on the array from location $\theta$, is written as

$$a(f_1, \theta) \triangleq \left[ e^{-i2\pi f_1\tau_1(\theta)} \ldots e^{-i2\pi f_1\tau_M(\theta)} \right]^T.$$ (2.2)

The $k$th snapshot from the $l$th FFT bin is written as

$$x_{k,f_l} \triangleq \left[ x_{1,k}(f_l) \ldots x_{M,k}(f_l) \right]^T = a_0(f_l, \theta_0)s_{0,k}(f_l) + n_{k,f_l},$$ (2.3)

where the zero-mean complex Gaussian noise vector $n_{k,f_l}$ is defined similarly to $x_{k,f_l}$. The cross-spectral density matrix (CSDM) at frequency $f_l$ is given by $R_{f_l} = E \left\{ x_{k,f_l}x_{k,f_l}^H \right\} = \sigma_0^2 a_0(f_l, \theta_0)a_0^H(f_l, \theta_0) + N_{f_l}$, where $\sigma_0^2 a_0 = E \left\{ |s_{0,k}(f_l)|^2 \right\}$ is the desired signal power at frequency $f_l$ and $N_{f_l} = E \left\{ n_{k,f_l}n_{k,f_l}^H \right\}$ is the noise-plus-interference covariance. In practice, the true CSDM is not available and $R_{f_l}$ is replaced by the sample CSDM estimate

$$\hat{R}_{f_l} = \frac{1}{K} \sum_{k=1}^{K} x_{k,f_l}x_{k,f_l}^H.$$ (2.4)
3. Fast Conjugate Gradient-Based RDRCB

Here, we summarise the conjugate-gradient based RDRCB, presented in [Somasundaram et al. 2013, Somasundaram et al. (submitted)], describing how to obtain the weights for the uth beam and ith FFT bin. In the following, \( \hat{a} \) denotes the assumed ASV, given by substituting the assumed steer direction \( \bar{\theta}_b \) for the \( b \)th beam into (2.2), yielding \( \hat{a} = a(f_i, \bar{\theta}_b) \). Following the RCB approach, the true ASV is assumed to belong to an \( M \)-dimensional element-space ellipsoid, denoted \( \mathcal{E}_M(\hat{a}, \mathbf{E}) \), given by

\[
\mathcal{E}_M(\hat{a}, \mathbf{E}) = \{ \mathbf{a} \in \mathbb{C}^M \mid \mathbf{a} - \hat{\mathbf{a}}^\mathsf{H} \mathbf{E} (\mathbf{a} - \hat{\mathbf{a}}) \leq 1 \},
\]

with centre \( \hat{\mathbf{a}} \) and axes described by \( \mathbf{E} \). The Hermitian symmetric matrix \( \mathbf{E} \), herein assumed positive definite, is also dependent on frequency and beam. Further, in the following we set \( \hat{\mathbf{R}}_x = \hat{\mathbf{R}}_h \) in (2.4).

The CG algorithm [Hestenes & Stiefel 1952, Dietl (2001)] is used to expand an \( N \)-dimensional Krylov basis, by setting \( d_1 = \mathbf{a} \) and \( r_1 = -\hat{\mathbf{a}} \), and for \( i = 1, \ldots, N-1 \), calculating

\[
\begin{align*}
\alpha_i &= -\frac{d_i^\mathsf{H} r_i}{d_i^\mathsf{H} \hat{\mathbf{R}}_x d_i}, \\
r_{i+1} &= r_i + \alpha_i \hat{\mathbf{R}}_x d_i, \\
\beta_i &= \frac{d_{i+1}^\mathsf{H} \hat{\mathbf{R}}_x r_{i+1}}{d_i^\mathsf{H} \hat{\mathbf{R}}_x d_i}, \\
d_{i+1} &= -r_{i+1} + \beta_i d_i,
\end{align*}
\]

where \( \mathbf{D} = [d_1 \ldots d_N] \) is the CG-based Krylov-subspace basis. In practice, \( N \) is unknown \textit{a priori} and therefore, the following stopping rule, first proposed in [Honig & Goldstein 2002] for the MSWF, is used to select \( N \)

\[
N = \max \left\{ n : \frac{\| \Pi_{\mathbf{D}_{n-1}}^\perp \hat{\mathbf{R}}_x^2 \mathbf{a} \|^2}{\| \mathbf{R}_x^2 \mathbf{a} \|^2} < \delta \right\},
\]

where \( \delta \) is a small positive constant and \( \Pi_{\mathbf{D}_{n-1}} = I - \mathbf{D}_{n-1} [\mathbf{D}_{n-1}^\mathsf{H} \mathbf{D}_{n-1}]^{-1} \mathbf{D}_{n-1}^\mathsf{H} \), with \( \mathbf{D}_{n-1} = [d_1 \ldots d_{n-1}] \). Methods for efficiently implementing (3.3) with the CG method are described in [Somasundaram et al. (submitted)].

In reduced-dimension robust Capon beamforming, a dimension reducing transform \( \mathbf{D} \) is used to reduce the dimensions of \( \mathcal{E}_M(\hat{a}, \mathbf{E}) \), to produce a reduced-dimension ellipsoid \( \mathcal{E}_N(\hat{\mathbf{b}}, \mathbf{F}) \), where \( \hat{\mathbf{b}} = \mathbf{D}^\mathsf{H} \hat{\mathbf{a}} \) and (for non-degenerate ellipsoids) \( \mathbf{F} = [\mathbf{D}^\mathsf{H} \mathbf{E}^{-1} \mathbf{D}]^{-1} \), and to reduce the dimensions of the data, by producing \( \hat{\mathbf{R}}_y = \mathbf{D}^\mathsf{H} \hat{\mathbf{R}}_x \mathbf{D} \). Then, the RDRCB optimization problem becomes [Somasundaram 2011, Somasundaram et al. 2013, Somasundaram et al. (submitted)]

\[
\min_{\mathbf{b}} \mathbf{b}^\mathsf{H} \hat{\mathbf{R}}_y^{-1} \mathbf{b} \text{ s.t. } \begin{bmatrix} \mathbf{b} - \hat{\mathbf{b}} \end{bmatrix}^\mathsf{H} \mathbf{F} \begin{bmatrix} \mathbf{b} - \hat{\mathbf{b}} \end{bmatrix} \leq 1,
\]

which, in general, requires two eigenvalue decompositions (EVDs) to solve. However, when \( \mathbf{D} \) is formed using the CG algorithm, \( \hat{\mathbf{R}}_y = \mathbf{D}^\mathsf{H} \hat{\mathbf{R}}_x \mathbf{D} = \mathbf{A}_{\text{CG}} \), where \( \mathbf{A}_{\text{CG}} \) is a diagonal matrix given by \( \mathbf{A}_{\text{CG}} = \text{diag} \{ [d_1^\mathsf{H} \hat{\mathbf{R}}_x d_1 \ldots d_N^\mathsf{H} \hat{\mathbf{R}}_x d_N] \} \). Therefore, \( \hat{\mathbf{R}}_y^{-1} = \mathbf{A}_{\text{CG}}^{-1} \), so that (3.4) can be written as

\[
\min_{\mathbf{b}} \mathbf{b}^\mathsf{H} \mathbf{A}_{\text{CG}}^{-1} \mathbf{b} \text{ s.t. } \begin{bmatrix} \mathbf{b} - \hat{\mathbf{b}} \end{bmatrix}^\mathsf{H} \mathbf{F} \begin{bmatrix} \mathbf{b} - \hat{\mathbf{b}} \end{bmatrix} \leq 1.
\]
EXPERIMENTAL RESULTS

4. Experimental Results

Here, we evaluate the CG-RDRCB, exploiting (tight) spherical and non-degenerate minimum volume ellipsoidal (NDMVE) uncertainty sets, the standard CG-MVDR and the DAS beamformers on data recorded on a $M \approx 300$ element array consisting of 8 rows of sensors (see [Somasundaram 2011] for further details). For the horizontal elevation,
96 equally spaced azimuth beams were formed. Using the methods in [Somasundaram et al. 2012], we design NDMVE sets, whose error sphere radii are set to $\hat{\epsilon} = 10$, and tight spherical uncertainty sets, based on the expected angle-of-arrival errors given the spacing of the beams. We refer the reader to [Somasundaram et al. 2012] for further details. Due to the dynamic environment in which the array is operating, only $K = 80$ snapshots are available for CSDM estimation. The data contains a strong controlled acoustic source on which we can check the ability of the beamformer to protect against SOI cancellation when pointing towards the source and its ability to null the source out when pointing away from it. Here, we only examine results from the frequency bin at which the array elements are approximately half-wavelength spaced. Fig. 1 illustrates the bearing time records, which indicate the strong acoustic source moves through most azimuth angles. The high DAS sidelobes, which can be seen clearly in Fig. 1 (a), are masking weak sources around azimuth angle indices 15, 50, 65 and 80, which can be seen when using the adaptive CG-based algorithms. The robust CG-Spherical and CG-NDMVE algorithms protect against SOI cancellation for beams pointing at the controlled acoustic source (as their powers converge to the DAS beamformer), whereas CG-MVDR does not.

REFERENCES


