A COMPARISON OF INITIALIZATION SCHEMES FOR BLIND ADAPTIVE BEAMFORMING

Thomas E. Biedka

Raytheon E-Systems, Inc.
Airborne Systems Division
Greenville TX 75403

Mobile and Portable Radio Research Group
Dept. of Electrical and Computer Engineering
Virginia Tech
Blacksburg VA 24061

ABSTRACT

Many blind adaptive beamforming algorithms require the selection of one or more non-zero initial weight vectors. Proper selection of the initial weight vectors can speed algorithm convergence and help ensure convergence to the desired solutions. Three alternative initialization approaches are compared here, all of which depend only on second order statistics of the observed data. These methods are based on Gram-Schmidt orthogonalization, eigendecomposition, and QR decomposition of the observed data covariance matrix. We show through computer simulation that the eigendecomposition approach yields the best performance.

1. INTRODUCTION

The problem of separating multiple cochannel communication signals which are received at an antenna array has received considerable interest. An important application is in smart antennas for mobile wireless communications [1]. One approach for separating the received signals is through blind adaptive beamforming. Blind adaptive algorithms may be defined as those adaptive algorithms which do not require the presence of a known training signal. In many cases the use of a blind adaptive beamforming algorithm requires the selection of a set of non-zero initial weight vectors. Examples of algorithms which require non-zero initial weight vectors include those based on constant modulus properties (e.g., [2, 3, 4, 5, 6, 7]) and those based on finite alphabet properties (e.g., [8, 9]). Proper selection of the initial weight vectors is crucial since improper selection can slow or prevent algorithm convergence.

One possible approach for selecting the initial weight vectors is to perform angle of arrival (AOA) estimation, and use this information to form the initial weight vectors. However, this process is computationally intensive, and requires that the antenna array response be known or experimentally calibrated. Array calibration is in practice very difficult to obtain and maintain. One of the main advantages of blind adaptive beamforming is that it does not require AOA information. For these reasons we do not consider initialization schemes which require AOA estimation or knowledge of the array response.

This work was supported by Rome Laboratories RL/IRAA and by Raytheon E-Systems, Airborne Systems Division.

The initial weight vectors considered here are derived directly from the sample covariance matrix of the observed data. Since these weight vectors do not depend on any properties of the received signals, such as modulation format, they can be used in many different adaptive algorithms. The initialization methods considered here are based on

1) Gram-Schmidt orthogonalization,
2) Eigendecompositon,
3) QR decomposition.

All of these methods yield orthogonal, or nearly orthogonal, output signals. This helps to ensure that independent solutions are found by each beamformer weight vector. These methods are described in more detail in the following sections. Following this desciption, the results of Monte Carlo simulations are presented to compare performance. Performance is examined both as a function of power spread and angular spread of the received signals.

2. GRAM-SCHMIDT INITIALIZATION

This approach uses a standard Gram-Schmidt algorithm to generate a set of weight vectors which orthogonalize the data. These weight vectors are then used as the initial weight vectors for the adaptive algorithm. It should be noted that this set of weight vectors is not unique.

Suppose that we wish to generate a set of L weight vectors that will orthogonalize the received data by using a Gram-Schmidt procedure. Denote the set of L initial weight vectors by the $M \times L$ matrix $\mathbf{W}$, where M is the number of elements in the array. As a first step we set the main diagonal of $\mathbf{W}$ equal to L ones, while the remainder of $\mathbf{W}$ is set to zeros. That is, the first column of $\mathbf{W}$ is set to $\mathbf{w}_1 = [1\ 0\ 0\ \cdots 0]^T$, the second column is set to $\mathbf{w}_2 = [0\ 1\ 0\ \cdots 0]^T$, etc. The first Gram-Schmidt weight vector, which we will denote by $\hat{\mathbf{w}}_1$, is equal to $\mathbf{w}_1$. The second Gram-Schmidt weight vector $\hat{\mathbf{w}}_2$ is found by

$$\hat{\mathbf{w}}_2 = \mathbf{w}_2 - \alpha_{1,2}\hat{\mathbf{w}}_1 \tag{1}$$

where

$$\alpha_{i,j} = \frac{\hat{\mathbf{w}}_i^H \mathbf{R}_{\mathrm{xx}} \mathbf{w}_j}{\hat{\mathbf{w}}_i^H \mathbf{R}_{\mathrm{xx}} \hat{\mathbf{w}}_i} \tag{2}$$

and $\mathbf{R}_{\mathrm{xx}}$ is the observed data sample covariance matrix. In general, the kth Gram-Schmidt weight vector is given by

$$\hat{\mathbf{w}}_k = \mathbf{w}_k - \sum_{i=1}^{k-1} \alpha_{i,k} \hat{\mathbf{w}}_i. \tag{3}$$

One advantage of the Gram-Schmidt approach over the eigen- or QR-decomposition approaches is lower computational complexity.

3. EIGENVECTOR AND QR INITIALIZATION

The eigenvector initialization uses the L most dominant eigenvectors of the observed data covariance matrix as the L initial weight vectors. We first show that an optimal beamformer weight vector lies in the signal subspace of the observed data covariance matrix. The weight vector $\mathbf{w}_{opt}$ that extracts a desired signal with maximum output Signal to Interference and Noise Ratio (SINR) is given by

$$\mathbf{w}_{opt} \propto \mathbf{R}_{\mathrm{xx}}^{-1}\mathbf{a}, \tag{4}$$

where $\mathbf{a}$ is the array response vector of the desired signal and $\mathbf{R}_{\mathrm{xx}}$ is the observed data correlation matrix. In the case where the background noise is white, and there are fewer incident signals L than sensors M, the observed data correlation matrix may be expressed as

$$\mathbf{R}_{\mathrm{xx}} = \mathbf{U}_{\mathrm{S}}\mathbf{\Sigma}_{\mathrm{S}}\mathbf{U}_{\mathrm{S}}^H + \sigma_{\mathrm{N}}\mathbf{U}_{\mathrm{N}}\mathbf{U}_{\mathrm{N}}^H \tag{5}$$

where $\mathbf{U}_{\mathrm{S}}$ is an $M \times L$ matrix of signal subspace eigenvectors, $\mathbf{\Sigma}_{\mathrm{S}}$ is an $L \times L$ diagonal matrix of the corresponding eigenvalues, $\mathbf{U}_{\mathrm{N}}$ is an $M \times (M - L)$ matrix of signal nullspace eigenvectors, and σ_{N} is the power of the background noise. Since the steering vector $\mathbf{a}$ lies in the signal subspace, $\mathbf{U}_{\mathrm{N}}^H\mathbf{a} = 0$, which in turn implies

$$\begin{aligned} \mathbf{R}_{\mathrm{xx}}^{-1}\mathbf{a} &= (\mathbf{U}_{\mathrm{S}}\mathbf{\Sigma}_{\mathrm{S}}^{-1}\mathbf{U}_{\mathrm{S}}^H + \frac{1}{\sigma_{\mathrm{N}}}\mathbf{U}_{\mathrm{N}}\mathbf{U}_{\mathrm{N}}^H)\mathbf{a} & (6) \\ &= \mathbf{U}_{\mathrm{S}}\mathbf{\Sigma}_{\mathrm{S}}^{-1}\mathbf{U}_{\mathrm{S}}^H\mathbf{a} & (7) \end{aligned}$$

Since we know that the adaptive weight vectors will lie in the signal subspace upon convergence, it makes sense for the initial weight vectors to lie in the signal subspace.

Another motivation is that the eigenvectors themselves can be very good beamformer weight vectors under some circumstances. In particular, as the power spread between the received signals increases, the eigenvectors become better beamformer weight vectors. This behavior can be explained as follows. Consider an environment with two incident signals, with one signal much stronger than the other. The dominant eigenvector, which by definition is the beamformer weight vector that maximizes the output power, will be very similar to the steering vector of the stronger signal. Thus the output of the first eigen-beamformer will be dominated by the stronger signal. The second most dominant eigenvector maximizes the output power subject to the constraint that the output is orthogonal to the first output. Thus the second eigenvector will extract the weaker signal with fairly high output SINR. As the power spread becomes higher, the second eigenvector will yield higher output SINR. The eigenvectors also become better weight vectors as the angular separation increases. Other work on the performance of eigenvectors as beamformer weight vectors has been presented in [10]. A related area is the well established use of eigenvectors as an orthogonalizing pre-processor for adaptive beamforming [11].

If the number of incident signals is less than the number of elements, some of the weight vectors will lie in the signal nullspace. For this reason it is important to consider what happens if we use a signal nullspace eigenvector as an initial weight vector. A nullspace eigenvector will null all of the incident signals, so that the beamformer output contains only background noise. If a CMA is used to adapt the weight vector, the algorithm may stay in this noise capture state indefinitely. The noise capture properties of CMA have been examined by several authors [12, 13]. It has been shown that noise capture corresponds to a saddlepoint in the CMA cost function.

The QR initialization is based on a QR decomposition of the observed data covariance matrix. Using this decomposition the covariance matrix can be expressed as

$$\mathbf{R}_{\mathrm{xx}} = \mathbf{QR} \tag{8}$$

where the matrix $\mathbf{R}$ is upper triangular and the matrix $\mathbf{Q}$ is orthonormal. The columns of $\mathbf{Q}$ are used as the initial weight vectors. Since $\mathbf{Q}$ does not orthogonalize the data, we expect the performance to be worse than the eigenvector initialization. However, the columns of $\mathbf{Q}$ are sometimes very similar to the eigenvectors of $\mathbf{R}_{\mathrm{xx}}$, so the behavior of the QR and eigendecomposition methods may be similar.

4. OVERVIEW OF LSCMA

We now give a brief overview of the Least Squares Constant Modulus Algorithm (LSCMA) [2], which we use to compare the different initialization schemes. This algorithm extracts one incident signal. It will typically extract the strongest constant modulus signal received by the array if the initial weight vector is omnidirectional.

Let the $M \times 1$ complex vector $\mathbf{x}(n)$ represent the signals and noise received at an array of M antennas. We denote the initial weight vector by $\mathbf{w}_0$, where we have changed notation so that the subscript 'i' on $\mathbf{w}_i$ now denotes the iteration number of the algorithm. Given N samples of observed data, an initial signal estimate is formed using the initial weight vector via

$$y_0(n) = \mathbf{w}_0^H\mathbf{x}(n). \tag{9}$$

The initial signal estimate is then hard limited to yield

$$d_0(n) = \frac{y_0(n)}{|y_0(n)|}. \tag{10}$$

A new weight vector is formed according to

$$\mathbf{w}_1 = \mathbf{R}_{xx}^{-1}\mathbf{r}_{xd} \tag{11}$$

where $\mathbf{R}_{xx} = \left\langle \mathbf{x}(n)\,\mathbf{x}^H(n) \right\rangle$, $\mathbf{r}_{xd} = \langle \mathbf{x}(n)\, d_0^*(n) \rangle$, and $\langle \cdot \rangle$ denotes a time average over $0 \leq n \leq N-1$. The iteration described by (9),(10), and (11) is then repeated with $\mathbf{w}_0$ replaced by $\mathbf{w}_1$. This process is continued until either the change in the weight vector is smaller than some threshold,

$\Delta\sigma^2$ (dB)	Initial SINR (dB) GSO	QR	Eigen.
0	-2.6	0.0	-0.1
0.1	-2.5	0.1	0.2
0.25	-2.4	0.2	0.6
0.5	-2.1	0.4	1.3
1.	-1.7	1.0	2.5
2.	-0.8	1.9	4.9
3.	0.1	2.9	7.1
6.	2.3	5.9	12.0
9.	4.0	8.8	15.6

Table 1: Expected value of initial output SINR for the weaker signal as a function of power spread.

$\Delta\theta$	Optimal SINR (dB)	Initial SINR (dB) GSO	QR	Eigen.
0°	-1.0	-25.2	-25.1	$-\infty$
1°	6.8	-17.5	-8.8	-2.0
2°	12.1	-12.2	-0.6	0.0
3°	15.4	-9.1	0.5	0.6
5°	19.3	-5.4	0.8	1.0
10°	23.4	-1.7	1.0	2.5
20°	23.8	0.1	1.0	4.2
30°	24.0	0.6	1.0	24.0

Table 2: Expected value of initial output SINR for the weaker signal as a function of angular separation.

or until the envelope variance of the output signal is deemed sufficiently small. The convergence rate of this algorithm is investigated in [3, 4].

5. SIMULATION RESULTS

We now present the results of computer simulations to compare the relative performance of the different initialization strategies. For all simulations we use an 8 element uniform linear array with $\lambda/2$ spacing. The incident signals are FM with independent random low-pass Gaussian modulating waveforms. All incident signals have the same carrier offset of zero. The background noise is temporally and spatially white with complex Gaussian distribution. The power of the received signals is measured relative to the background noise and is expressed in terms of the Signal to White Noise Ratio (SWNR).

Before we present results from Monte Carlo simulations, it is important to examine the expected initial SINR of each approach. The expected initial SINR can be easily calculated by generating the ideal, i.e., infinite collect time, covariance matrix. We will first consider the initial SINR versus power spread. The environment contains two signals, with one signal incident from 0° (broadside), and the second signal incident from 10°. The SWNR of the second signal is fixed at 15 dB. The SWNR of the first signal is varied from 15 dB to 24 dB. Table 1 compares the output SINR achieved for the second (weaker) signal with the different initial weight vectors. The output SINR of the stronger signal is always higher than for the weaker signal. All three methods show increasing output SINR with increasing power spread. The eigenvector approach has the highest overall output SINR. If the power spread is 3 dB, the output SINR is quite high at 7.1 dB. The GSO approach has the lowest output SINR, being negative until the power separation reaches 3 dB. The QR method performs somewhat worse than the eigenvector method, as expected.

Table 2 compares the initial SINR versus angular spread. The environment again contains two incident signals, with the AOA of the first signal fixed at 0° with 16 dB SWNR. The AOA of second signal is varied from 0° to 30° while the SWNR is fixed at 15 dB SWNR. Note that the power spread is fixed at 1 dB. Results are again shown for the weaker signal. When the signals have the same AOA, the eigenvector method yields an output SINR of 0, or $-\infty$ dB, since the second most dominant eigenvector is orthogonal to the steering vector of the incident signals. However, even at very small angular separation the eigenvector method yields dramatically higher initial SINR than the GSO method. Once again, the QR method performs slightly worse than the eigenvector method.

While it is important to examine the initial output SINR, the results presented so far do not include the effects of finite collect time. In addition, the initial output SINR is not sufficient to determine the solutions that LSCMA will converge to, particularly for low initial SINR. The following approach is used to determine the effectiveness of each initialization strategy. The goal is to extract all the incident signals. We first generate a realization of the received data. We then compute a set of initial weight vectors, which is dependent on the sample covariance matrix of the data. The number of initial weight vectors is equal to the number of incident signals. For all LSCMA simulations we use a block size of 100 samples. The LSCMA is then applied to the same data using these different initial weight vectors. Each set of LSCMA iterations is run independently, and no attempt is made to force each weight vector to a different solution. After 20 iterations, which is more than sufficient for LSCMA to converge, the solutions found are compared. If different solutions have been found, then the initialization has performed correctly. This procedure is repeated for each different initialization method for 1000 independent trials.

We first consider performance versus power spread. The environment is identical to the environment used to obtain the results in Table 1. The results in Table 3 show the likelihood that LSCMA will extract both signals, i.e., that the LSCMA weight vectors will converge to different solutions. As can be seen, the eigenvector method performs very well. It causes LSCMA to converge to the two desired solutions with as little as 0.5 dB of power separation. Consulting Table 1, we see that the expected initial output SINR of the eigenvector is just 1.3 dB, but this is sufficient to ensure that the second LSCMA weight vector extracts the weaker signal. The QR method performs only slightly worse than the eigendecomposition method, while the GSO method performs poorly in comparison.

Similar conclusions can be drawn from examining the

$\Delta\sigma^2$ (dB)	GSO	QR	Eigen.
0	.542	.685	0.653
0.1	.555	.707	0.820
0.25	.564	.870	0.981
0.5	.654	.982	1.
1	.766	1.	1.
2	.921	1.	1.
3	.970	1.	1.

Table 3: Fraction of trials in which LSCMA weight vectors extracted different signals versus received power difference.

$\Delta\theta$	GSO	QR	Eigen.
1°	.491	.528	.739
2°	.510	.839	.935
3°	.527	.958	.981
4°	.590	.997	.997
5°	.615	1.	0.999
10°	.766	1.	1.

Table 4: Fraction of trials in which LSCMA weight vectors extracted different signals versus angular separation.

likelihood of LSCMA convergence as a function of angular separation. The signal environment is identical to that used in Table 2. The eigenvector method performs best overall, and the QR method performs almost as well. The GSO method does not perform well.

Because of space limitations, results obtained in other signal environments can not be included here. However, we have observed that in the three signal environment the power spread must be on the order of 6 dB for the eigenvector method to cause LSCMA to extract all three signals. This is a much larger required power spread than in the two signal environment. Both the QR and GSO methods perform much worse than the eigenvector method in the three signal environment. This indicates that the eigenvector method should be the preferred choice.

6. CONCLUSIONS

We have compared the performance of several different initialization techniques for blind adaptive beamforming. The results of Monte Carlo simulation indicate that the dominant eigenvectors of the observed data covariance matrix yield good initial weight vectors. The eigenvectors encourage the algorithm to extract all the incident signals because these weights orthogonalize the data, and because the eigenvectors yield reasonable initial SINR in many cases.

The problem of initialization of blind adaptive equalizers is not considered here, but the eigenvector initialization might prove useful for this problem. Currently it is common to initialize a blind equalizer by setting the so-called 'center-tap' equal to a constant, while leaving the other taps set to zero. It is possible that eigenvector initialization may speed blind equalizer convergence, but the added computational complexity may not be warranted.

7. REFERENCES

[1] L.C. Godara, "Applications of antenna arrays to mobile communications, part I: Performance improvement, feasibility, and system communications", *Proc. of the IEEE*, vol. 85, no. 7, pp. 1031–1060, July 1997.

[2] B. G. Agee, "The least-squares CMA: A new technique for rapid correction of constant modulus signals", *Proc. of Inter. Conf. on Acous., Speech, and Signal Process.*, pp. 953–956, April 1986.

[3] T. E. Biedka, W. H. Tranter, and J. H. Reed, "Convergence analysis of the least squares constant modulus algorithm", *Proc. of the Asilomar Conf. on Signals, Systems, and Computers*, pp. 541–545, Nov. 1996.

[4] T. E. Biedka, W. H. Tranter, and J. H. Reed, "Convergence analysis of the least squares constant modulus algorithm in interference cancellation applications", *Submitted to IEEE Trans. on Commun.*, Sept. 1997.

[5] B. G. Agee, "Blind separation and capture of communication signals using a multitarget constant modulus beamformer", *Proc. of IEEE Military Commun. Conf.*, pp. 340–346, May 1989.

[6] J. J. Shynk and R. P. Gooch, "Convergence properties of the multistage CMA adaptive beamformer", *Proc. of the Asilomar Conf. on Signals, Systems, and Computers*, pp. 622–626, Nov. 1993.

[7] I. Parra, G. Xu, and H. Lui, "A least squares projective constant modulus approach", *Personal Indoor Mobile Radio Conference*, pp. 673–676, 1995.

[8] A.L. Swindlehurst, S. Daas, and J. Yang, "Analysis of a decision directed beamformer", *IEEE Trans. on Signal Processing.*, vol. 43, no. 12, December 1995.

[9] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array: Part i. algorithms", *IEEE Trans. on Signal Processing.*, vol. 44, no. 5, May 1996.

[10] M.A. Mow and M.F. Kahn, "A two-sensor digital power inversion array", *Proc. of the Asilomar Conf. on Signals, Systems, and Computers*, pp. 750–754, October 1995.

[11] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*, John Wiley and Sons, NY, 1980.

[12] J. R. Treichler and M. G. Larimore, "The noise capture properties of the constant modulus algorithm", *Proc. of Inter. Conf. on Acous., Speech, and Signal Process.*, pp. 1165–1168, August 1985.

[13] B. G. Agee, "Convergent behavior of modulus-restoring adaptive arrays in gaussian interference environments", *Proc. of the Asilomar Conf. on Signals, Systems, and Computers*, pp. 818–822, Dec. 1988.