ON LOW RANK MVDR BEAMFORMING USING THE CONJUGATE GRADIENT ALGORITHM

Ernesto L. Santos and Michael D. Zoltowski

School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 Email:{santose, mikedz}@ecn.purdue.edu

ABSTRACT

In array applications an important task is adaptive beamforming. The Minimum Variance Distortionless Response (MVDR) beamformer can only be computed if the true spatial correlation matrix is available. In practice, the correlation matrix has to be estimated from the arriving signals, and in some cases there are only a small number of samples (snapshots) available. When the number of snapshots is small, the MVDR beamformer is no longer optimal, and a low rank MVDR solution can provide a higher SINR. In this work we will analyze two methods of finding low rank solutions: Steering Independent Conjugate Gradient (SI-CG) and Steering Dependent Conjugate Gradient (SD-CG). We will also propose a simplified expression to compute the arriving power from any given direction.

1. INTRODUCTION

The MVDR beamformer is derived assuming that the true spatial correlation matrix is available. In practice, only an estimate of the correlation matrix is available. In some applications, there is only a limited number of snapshots available for estimating the spatial correlation matrix. In these cases, a low rank solution of the scaled Weiner-Hopf equations can yield a higher SINR than the full-rank MVDR beamformer. There are several methods of finding low rank MVDR beamformers, such as the Principal Components Inverse (PCI) algorithm [1] and the Multi-Stage Wiener Filter (MWF) of Goldstein et. al. [2]. In [3], Guido et. al. proved an equivalence between the MWF and the Conjugate Gradient (CG) Algorithm.

Dependent Adaptive Beamforming (ABF). The scenario assumes the formation of multiple adaptive beams, each pointed to a different "look" direction. In Steering-Dependent ABF, a Generalized Sidelobe Canceler (GSC) is formed for each "look" direction. Mathematically, the GSC serves to convert

the constrained MVDR optimization problem to an unconstrained optimization problem, thereby enforcing a-priori the unity gain constraint in the "look" direction. In contrast, in Steering-Independent ABF, a scaled version of the Wiener-Hopf equations is solved, and the unity gain constraint is enforced a-posteriori through simple scaling of the resulting ABF weight vector.

Implementation of a GSC for each "look" direction requires the construction and application to the data of a blocking matrix for each "look" direction. The attendant computational complexity is quite substantial. In this paper, we prove a very important and somewhat surprising result: the low-rank beamformer obtained with Steering Dependent Conjugate Gradients (SD-CG) is exactly the same as the low-rank beamformer obtained with Steering Independent Conjugate Gradients (SI-CG). This is an important result because Weippert et. al. [4] showed that a SD-CG significantly outperformed a Steering Independent version of PCI. Thus, our result dictates that we can obtain the performance of SD-CG without having to form blocking matrices for each "look" direction.

2. LOW RANK MVDR SOLUTIONS

2.1. Steering-Independent Beamforming

The MVDR beamformer is found by solving the constrained optimization problem:

min
$$S_{xx}(\theta) = \mathbf{w}^{H}(\theta) \mathbf{R} \mathbf{w}(\theta)$$

 $\mathbf{w}(\theta) \qquad s.t. \ \mathbf{w}^{H}(\theta) \mathbf{s}(\theta) = 1$ (1)

where $s(\theta)$ is the steering vector for a direction θ , **R** is the tween Steering-Independent Adaptive Beamforming and Steering-M is the number of elements in the array. $S_{xx}(\theta)$ is a measure of power arriving from the direction θ . In order to keep expressions simple we will omit the direction θ from both $\mathbf{w}(\theta)$ and $\mathbf{s}(\theta)$ in the following equations. However, it should be kept in mind that w and s are associated with a particular "look" direction θ . Using the Method of Lagrange Multipliers and taking the gradient of the augmented objective function dictates that

Supported by Offi ce of Naval Research grant N00014-03-1-0077.

the solution to the constrained optimization problem in (1) may be computed as the solution to

$$\mathbf{R}\mathbf{w} = \lambda \mathbf{s},\tag{2}$$

where the Lagrange multiplier serves to satisfy the unity gain constraint in (1). In Steering-Independent CG, the CG algorithm is used to solve $\mathbf{Rw}' = \mathbf{s}$; the solution obtained at step r of CG is denoted $\mathbf{w}^{'(r)}$. The unity gain constraint is satisfied a-posteriori by scaling $\mathbf{w}^{'(r)}$ according to

$$\mathbf{w}^{(r)} = \mathbf{w}^{\prime(r)} / (\mathbf{s}^H \mathbf{w}^{\prime(r)}) \tag{3}$$

We refer to this procedure of finding a low rank MVDR solution as SI-CG. This procedure solves $\mathbf{Rw'} = \mathbf{s}$ without satisfying the unity gain constraint a-priori. When CG is terminated, scaling is used a-posteriori to satisfy the constraint in (1). In contrast, SD-CG (developed in the next section), enforces the unit gain constraint a-priori through the use of a blocking matrix.

Dietl et. al. [3] observed that at each step r the output vector of the CG algorithm, $\mathbf{w}^{'(r)}$, minimizes the error metric

$$e(\mathbf{w}^{\prime(r)}) = (\mathbf{R}^{-1}\mathbf{s} - \mathbf{w}^{\prime(r)})^{H}\mathbf{R}(\mathbf{R}^{-1}\mathbf{s} - \mathbf{w}^{\prime(r)})$$
(4)

for a $\mathbf{w}^{'(r)}$ constrained to the Krylov subspace $\mathcal{K}_{si}^{(r)}$ defined by the column space of $\mathbf{K}_{si}^{(r)}$, where

$$\mathbf{K}_{si}^{(r)} = \left[\mathbf{s} \stackrel{:}{\cdot} \mathbf{Rs} \stackrel{:}{\cdot} \dots \stackrel{:}{\cdot} \mathbf{R}^{r-1} \mathbf{s} \right]$$
(5)

Since after r steps the solution is constrained to a r dimensional subspace; it is also referred as a rank r solution.

2.2. Steering-Dependent Beamforming

We use a blocking matrix $\mathbf{B}(\theta)$, which columns have unit norm and are made to be the orthogonal complement of the steering vector $\mathbf{s}(\theta)$. Thus we can write the desired beamformer as the sum of the steering vector and a linear combination of orthornormal vectors, which are the columns of $\mathbf{B}(\theta)$,

$$\mathbf{w}(\theta) = \mathbf{s}(\theta) + \mathbf{B}(\theta)\mathbf{u}(\theta) \tag{6}$$

Assuming, without loss of generality, that $\mathbf{s}(\theta)$ has unit norm for each look direction θ , $\mathbf{s}(\theta)^H \mathbf{s}(\theta) = 1$. To keep expression simple we, again, omit the direction θ . Substituting (6) into equation (1), the optimization problem becomes unconstrained, since the constraint $\mathbf{s}^H \mathbf{w} = 1$ is always satisfied. The problem can be re-stated as:

min
$$(\mathbf{s} + \mathbf{B}\mathbf{u})^H \mathbf{R}(\mathbf{s} + \mathbf{B}\mathbf{u})$$

 \mathbf{u} (7)

Taking the gradient of (7) yields the optimal value of \mathbf{u} as the solution to

$$(\mathbf{B}^H \mathbf{R} \mathbf{B}) \mathbf{u} = -\mathbf{B}^H \mathbf{R} \mathbf{s} \tag{8}$$

Similarly to SI-CG we use the CG algorithm for low rank solutions $\mathbf{u}^{(r)}$ to (8). Substituting $\mathbf{u}^{(r)}$ for \mathbf{u} into (6) we obtain a low rank beamformer \mathbf{w} which is referred as $\mathbf{w}^{(r+1)}$. We added one to the rank because the steering vector is already in the subspace. We refer to this procedure as SD-CG. Using the CG algorithm $\mathbf{u}^{(r)}$ is restricted to the Krylov subspace $\mathcal{K}_{u}^{(r)}$ defined by the column space of $\mathbf{K}_{u}^{(r)}$, where

$$\mathbf{K}_{u}^{(r)} = [\mathbf{B}^{H}\mathbf{Rs} \stackrel{:}{:} (\mathbf{B}^{H}\mathbf{RB})\mathbf{B}^{H}\mathbf{Rs} \stackrel{:}{:} \dots \\ (\mathbf{B}^{H}\mathbf{RB})^{r-1}\mathbf{B}^{H}\mathbf{Rs}]$$
(9)

Using equations (6) and (9), the subspace that $\mathbf{w}^{(r)}$ is restricted to is $\mathcal{R}(\mathbf{K}_{sd}^{(r)})$, where the operator \mathcal{R} denotes the range space of the matrix and

$$\mathbf{K}_{sd}^{(r)} = \left[\mathbf{s} \stackrel{:}{:} \mathbf{B}\mathbf{K}_{u}^{(r-1)} \right]$$
(10)

After some algebraic manipulation, we have

$$\mathbf{K}_{sd}^{(r)} = [\mathbf{s} : (\mathbf{B}\mathbf{B}^{H}\mathbf{R})\mathbf{s} : \dots : (\mathbf{B}\mathbf{B}^{H}\mathbf{R})^{r-1}\mathbf{s}] \quad (11)$$

3. EQUIVALENCE BETWEEN SI-CG AND SD-CG

In this section we prove that SI-CG and SD-CG yield the same beamformer for a same rank solution. First we prove that the SI-CG and SD-CG rank *r* beamformers are restricted to the same subspace; i.e., we show that the range of the matrices $\mathbf{K}_{si}^{(r)}$ and $\mathbf{K}_{sd}^{(r)}$ from equations (5), (11) respectively are the same. Then we derive closed-form expressions for both SI-CG and SD-CG beamformers and show that they are the same.

3.1. Same Subspaces

When a vector is pre-multiplied by the matrix \mathbf{BB}^{H} the component in the direction of s is removed and the components orthogonal to s are kept unchanged. Thus

$$\operatorname{span}\{\mathbf{s}, \mathbf{Rs}\} = \operatorname{span}\{\mathbf{s}, \mathbf{BB}^{H}\mathbf{Rs}\} \Rightarrow \mathcal{K}_{si}^{(2)} = \mathcal{K}_{sd}^{(2)} \quad (12)$$

That is, the beamformers from a rank two SD-CG solution and a rank two SI-CG solution are restricted to the same subspace.

The following theorem states that if the SD-CG and SI-CG Krylov subspaces of dimension r are the same, then the Krylov subspaces of dimension (r+1) will also be the same. To simplify notation, we define $\mathbf{C} = \mathbf{BB}^{H}\mathbf{R}$. Theorem

If span{s, Cs, ..., C^rs} = span{s, Rs, ..., R^rs}, then span{s, Cs, ..., C^{r+1}s} = span{s, Rs, ..., R^{r+1}s} Proof

- i) span{s, $Cs, \ldots, C^{r+1}s$ } = span{s, $Rs, \ldots, R^rs, C^{r+1}s$ }, by hypothesis
- ii) $\mathbf{C}^{r+1}\mathbf{s} = \mathbf{C}(\mathbf{C}^r\mathbf{s}) = \mathbf{C}(\sum_{i=0}^r \alpha_i \mathbf{R}^i \mathbf{s}) = \sum_{i=0}^r \alpha_i \mathbf{C} \mathbf{R}^i \mathbf{s} = \mathbf{B} \mathbf{B}^H \sum_{i=0}^r \alpha_i \mathbf{R}^{i+1} \mathbf{s}$
- iii) span{s, Rs, ..., $\mathbf{R}^r \mathbf{s}, \mathbf{C}^{r+1}\mathbf{s}$ } = span{s, Rs, ..., $\mathbf{R}^r \mathbf{s}, \mathbf{BB}^H \sum_{i=0}^r \alpha_i \mathbf{R}^{i+1}\mathbf{s}$ } = span{s, Rs, ..., $\mathbf{R}^{r+1}\mathbf{s}$ }, because \mathbf{BB}^H only removes the component in the direction of s from the vectors $\mathbf{R}^{i+1}\mathbf{s}$, for i = 0 to r, and s is already in the set as well as $\mathbf{R}^{i+1}\mathbf{s}$, for i = 0 to r - 1.

From i) and iii) span{s, $Cs, \ldots, C^{r+1}s$ } = span{s, $Rs, \ldots, R^{r+1}s$ } Q.E.D.

Since we showed that $\mathcal{K}_{si}^{(2)} = \mathcal{K}_{sd}^{(2)}$, we have proved by induction that for $r \geq 2$: span $\{\mathbf{s}, \mathbf{Rs}, \mathbf{R}^2 \mathbf{s}, \dots, \mathbf{R}^{r-1} \mathbf{s}\} =$ span $\{\mathbf{s}, (\mathbf{BB}^H \mathbf{R}) \mathbf{s}, (\mathbf{BB}^H \mathbf{R})^2 \mathbf{s}, \dots, (\mathbf{BB}^H \mathbf{R})^{r-1} \mathbf{s}\}$ Thus, the subspace the rank r SI-CG beamformer is constrained to lie in is the same as the subspace that the rank r SD-CG beamformer is constrained to lie in.

3.2. Closed expressions for w_{si} and w_{sd}

3.2.1. SI-CG

Since we have proved that the rank r beamformers for SI-CG and SD-CG lie within the same subspace, the unscaled (prior to scaling to satisfy the unity gain constraint) SI-CG beamformer $\mathbf{w}_{si}^{'(r)}$ may be expressed as

$$\mathbf{w}_{si}^{'(r)} = \mathbf{K}_{si}^{(r)}\mathbf{y} = \mathbf{K}_{sd}^{(r)}\mathbf{x}$$
(13)

At this point in the development, \mathbf{K}_{sd} is known, and we seek an expression for \mathbf{x} . For uniqueness of \mathbf{x} , we assume that the columns of $\mathbf{K}_{sd}^{(r)}$ are linearly independent. If they are not a linearly independent set, then, without loss of generality, we redefine $\mathbf{K}_{sd}^{(r)}$ by removing an appropriate number of columns to achieve independence while maintaining the same column space.

The superscript r is omitted in the following development for purposes of notational simplicity. \mathbf{w}_{si} minimizes the objective function in (4). Substituting the expression for \mathbf{w}_{si} in (13) into (4), and taking the gradient and setting it to zero yields

$$\mathbf{K}_{sd}^{H}\mathbf{R}\mathbf{K}_{sd}\mathbf{x} - \mathbf{K}_{sd}^{H}\mathbf{s} = \mathbf{0}$$

$$\Rightarrow \mathbf{x} = (\mathbf{K}_{sd}^{H}\mathbf{R}\mathbf{K}_{sd})^{-1}\mathbf{K}_{sd}^{H}\mathbf{s}$$
(14)

Substituting the expression for \mathbf{K}_{sd} in (10), the inverse of $\mathbf{K}_{sd}^H \mathbf{R} \mathbf{K}_{sd}$ may be expressed as

$$(\mathbf{K}_{sd}^{H}\mathbf{R}\mathbf{K}_{sd})^{-1} = \begin{bmatrix} \mathbf{s}^{H}\mathbf{R}\mathbf{s} & \mathbf{s}^{H}\mathbf{R}\mathbf{B}\mathbf{K}_{u} \\ \mathbf{K}_{u}^{H}\mathbf{B}\mathbf{R}\mathbf{s} & \mathbf{K}_{u}^{H}\mathbf{B}^{H}\mathbf{R}\mathbf{B}\mathbf{K}_{u} \end{bmatrix}^{-1}$$
(15)

In addition, it follows that

$$\mathbf{K}_{sd}^{H}\mathbf{s} = [1 \vdots \mathbf{0}]^{H} \tag{16}$$

Substituting (15) and (16) into (14) dictates that only the first column of the inverse matrix in (15) is needed. Invoking the block matrix inversion lemma, x may be computed as

$$\mathbf{x} = \beta \begin{bmatrix} 1 \\ -(\mathbf{K}_u^H \mathbf{B}^H \mathbf{R} \mathbf{B} \mathbf{K}_u)^{-1} \mathbf{K}_u^H \mathbf{B} \mathbf{R} \mathbf{s} \end{bmatrix}$$
(17)

where β is a multiplicative scalar that will be accounted for when the unity gain constraint is satisfied. Substituting (17) into (13), with \mathbf{K}_{sd} given by (10) yields

$$\mathbf{w}_{si}' = \beta(\mathbf{s} - \mathbf{B}\mathbf{K}_u(\mathbf{K}_u^H \mathbf{B}^H \mathbf{R}\mathbf{B}\mathbf{K}_u)^{-1}\mathbf{K}_u^H \mathbf{B}\mathbf{R}\mathbf{s}) \quad (18)$$

Finally, scaling the above to satisfy the unity gain constraint $\mathbf{w}_{si}^H \mathbf{s} = 1$ yields

$$\mathbf{w}_{si} = \mathbf{s} - \mathbf{B}\mathbf{K}_u (\mathbf{K}_u^H \mathbf{B}^H \mathbf{R} \mathbf{B} \mathbf{K}_u)^{-1} \mathbf{K}_u^H \mathbf{B} \mathbf{R} \mathbf{s}$$
(19)

3.2.2. SD-CG

The rank r SD-CG beamformer may be expressed as

$$\mathbf{w}_{sd} = \mathbf{s} + \mathbf{B}\mathbf{u} = \mathbf{s} + \mathbf{B}\mathbf{K}_u\mathbf{z} \tag{20}$$

Since **u** is determined via the CG algorithm, it is the vector that minimizes the objective function

$$e_2(\mathbf{u}) = \mathbf{u}^H \mathbf{B}^H \mathbf{R} \mathbf{B} \mathbf{u} + \mathbf{u}^H \mathbf{B}^H \mathbf{R} \mathbf{s} + \mathbf{s}^H \mathbf{R} \mathbf{B} \mathbf{u} + \text{const}$$
(21)

with u constrained to lie within the range space of the matrix \mathbf{K}_u in (9). Substituting $\mathbf{K}_u \mathbf{z}$ for u above dictates that z is the vector that minimizes

$$\mathbf{z}^{H}\mathbf{K}_{u}^{H}\mathbf{B}^{H}\mathbf{R}\mathbf{B}\mathbf{K}_{u}\mathbf{z} + \mathbf{z}^{H}\mathbf{K}_{u}^{H}\mathbf{B}^{H}\mathbf{R}\mathbf{s} + \mathbf{s}^{H}\mathbf{R}\mathbf{B}\mathbf{K}_{u}\mathbf{z} + \text{const}$$

Taking the gradient and setting it to zero yields

$$\mathbf{z} = -(\mathbf{K}_u^H \mathbf{B}^H \mathbf{R} \mathbf{B} \mathbf{K}_u)^{-1} \mathbf{K}_u^H \mathbf{B}^H \mathbf{R} \mathbf{s}$$
(22)

In turn, substituting z back into (20) yields

$$\mathbf{w}_{sd} = \mathbf{s} - \mathbf{B}\mathbf{K}_u (\mathbf{K}_u^H \mathbf{B}^H \mathbf{R} \mathbf{B} \mathbf{K}_u)^{-1} \mathbf{K}_u^H \mathbf{B}^H \mathbf{R} \mathbf{s} \quad (23)$$

Equations (19) and (23) are observed to be identical, thereby proving that the rank r SI-CG beamformer is equal to the rank r SD-CG beamformer.

4. COMPUTATION OF THE POWER SPECTRUM

A measure of the power arriving from a direction θ is given by the function to be minimized in (1). If the exact MVDR beamformer w from equation (2) is used, then the expression for S_{xx} in (1) can be simplified to $S_{xx} = 1/\mathbf{s}^H \mathbf{w}$. We show that this simplification is also valid when w is substituted by a low rank solution, $\mathbf{w}^{(r)}$ using the CG algorithm.

 $\mathbf{w}_{si}^{'(r)}$ is already known to be the vector in the *r* dimension Krylov subspace that minimizes the error function given in (4). From equation (13), $\mathbf{w}_{si}^{'(r)} = \mathbf{K}_{si}^{(r)}\mathbf{y}$, where $\mathbf{K}_{si}^{(r)}$ is given in (5). Since $\mathbf{w}_{si}^{'(r)}$ minimizes the error function (4), \mathbf{y} can be found by solving,

$$\min_{\mathbf{y}} (\mathbf{y}^H \mathbf{K}_{si}^{(r) H} - \mathbf{s}^H \mathbf{R}^{-1}) \mathbf{R} (\mathbf{K}_{si}^{(r)} \mathbf{y} - \mathbf{R}^{-1} \mathbf{s})$$
(24)

Taking the gradient and setting it to zero

$$\mathbf{K}_{si}^{(r)\ H}\mathbf{R}\mathbf{K}_{si}^{(r)}\mathbf{y} - \mathbf{K}_{si}^{(r)\ H}\mathbf{s} = \mathbf{0}$$
(25)

Solving the above equation for y, the rank r SI-CG beamformer is given by,

$$\mathbf{w}_{si}^{\prime(r)} = \mathbf{K}_{si}^{(r)} \mathbf{y} = \mathbf{K}_{si}^{(r)} (\mathbf{K}_{si}^{(r) \ H} \mathbf{R} \mathbf{K}_{si}^{(r)})^{-1} \mathbf{K}_{si}^{(r) \ H} \mathbf{s}$$
(26)

 $\mathbf{w}_{si}^{'(r)}$ is the unscaled low-rank solution to equation (2) obtained via SI-CG; the unity gain constraint is satisfied aposteriori via (3). After a fair amount of algebraic manipulations, not included here due to space limitations, we obtain the following dramatic simplification

$$S_{xx} = (\mathbf{w}_{si}^{'(r)\ H} \mathbf{R} \mathbf{w}_{si}^{'(r)}) / (|\mathbf{w}_{si}^{'(r)\ H} \mathbf{s}|^2) = (\mathbf{w}_{si}^{'(r)\ H} \mathbf{s})^{-1}$$
(27)

That is, the arriving power obtained with the rank r CG beamformer may be computed via the far right hand side of (27). With this result a considerable amount of computation is avoided. Computing $\mathbf{w}_{si}^{'(r)} \mathbf{R} \mathbf{w}_{si}^{'(r)}$ would require $M^2 + M$ multiplications, while $1/\mathbf{w}_{si}^{'(r)}$ is requires only M multiplications. This computational savings by a factor of roughly M, the number of sensors (which may be quite large), is obtained at each and every "look" direction.

5. SIMULATION RESULTS

A simulation was conducted with a uniform linear array (half-wavelength spacing) composed of 24 elements. In addition to a desired signal with an SNR of 10 dB, there were 16 interferers with SNR's ranging from 10 dB to 30 dB. The Gaussian noise was temporally and spatially white. For each Monte Carlo run, the sample correlation matrix was estimated from 24 snapshots. Fig. 1 plots the output SINR associated with the desired signal obtained with both a rank r SI-CG beamformer and a rank r SD-CG beamformer, as a function of the value of r. It is observed that

SI-CG and SD-CG yield the same SINR at each rank dimension, as expected since we have proven they yield the same beamformer. For purposes of comparison, the horizontal line indicates the optimal SINR that would be obtained with the MVDR beamformer computed from the true correlation matrix.

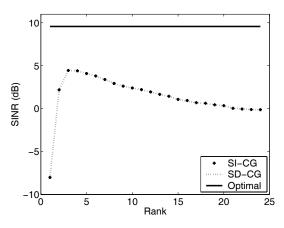


Fig. 1. SINR Performance: SI-CG vs. SD-CG

6. CONCLUSION

In this paper we proved that the low rank MVDR methods SD-CG and SI-CG yield the same beamformer. Thus, there is no reason to use SD-CG which is substantially more complex than SI-CG due to the formation and application of blocking matrices to the data, a different blocking matrix for each look direction. For SI-CG, we also developed a reduced complexity expression for computing the estimated power arriving from a given direction.

7. REFERENCES

- I. P. Kirsteins, D. W. Tufts "Adaptive detection using low rank approximation to a data matrix," *IEEE Trans. on Aerospace and Eletronic Systems*, vol. 30, no. 1, pp. 55-67, January 1994.
- [2] J. S. Goldstein, I. S. Reed and L. L. Scharf "A multistage representation of the wiener filter based on orthogonal projections," *IEEE Trans. on Information Theory*, vol. 44, no. 7, pp. 2943-2959, November 1988.
- [3] G. Dietl, M.D. Zoltowski and M. Joham, "Reducedrank equalization for EDGE via conjugate gradient implementation of multi-stage nested wiener filter," *IEEE Proc. of Vehi. Tech. Conf., VTC 2001*, vol. 3, pp. 1912-1916, October 2001.
- [4] M. Weippert, J. Hiemstra, J. S. Goldstein, V. Sabo, M. D. Zoltowski, I. Reed. "Signal-Dependent Reduced-Rank Multibeam Processing," *Radar 2003 - International Radar Conference*, Adelaide, Australia, 3-5 September 2003.