ATA memo #62

Automatic Interference Mitigation in Large N Correlator Systems Using an Eigenfilter

Jan 14 2004

Wilbert Lynn Urry

abstract

An excellent paper on the subject of interference mitigation by Amir Leshem and Alle-Jan Van Der Veen¹ proposed a technique to identify and extract the "spatial signature" of interference from observations in radio astronomy. The technique involves the short term measurement of the covariance matrix and the determination of its eigenvectors. I propose a method that is closely related to the method they describe but that can operate as an eigenfilter² on the data samples as they emerge from an interferometer. Using the ATA correlator architecture, the system would require only two multiply accumulate stages per antenna and some arithmetic to find the roots of a quadratic equation.

Imaging: Let v_i and v_j represent the complex samples emerging from antennas i and j at a single instant in time. Then any baseline may be calculated by

 $b_{ij} = v_i v_j^* \tag{1}$

If a column vector is formed out of all of the above antenna samples then a matrix of baselines can be formed from its conjugate square or outer product with itself.

Γ., Τ

$$\begin{bmatrix} v \end{bmatrix} = \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{N-1} \end{bmatrix}$$
(2)

$$B = \begin{bmatrix} b_{00} & b_{10} & \cdots & b_{N-1,0} \\ b_{01} & b_{11} & \cdots & b_{N-1,1} \\ \vdots & \vdots & \ddots & \vdots \\ b_{0,N-1} & b_{1,N-1} & \cdots & b_{N-1,N-1} \end{bmatrix} = \begin{bmatrix} v_0 \\ v_1 \\ \vdots \\ v_{N-1} \end{bmatrix} \begin{bmatrix} v_0^* & v_1^* & \cdots & v_{N-1}^* \end{bmatrix}$$
(3)

1

The baseline matrix B_i is Hermitian so that only (N+1)N/2 of the numbers are unique. The real parts of b_{ij} are symmetric about the diagonal and the imaginary parts are asymmetric. The order of the indices is important since $\angle b_{ij} = -\angle b_{ij}$.

Assume that D represents the discrete Fourier transform of the sample vector v resulting in the voltage image sample s. In matrix notation the discrete Fourier transform may be represented as:

 $[s(0)] = [D][v(0)] \tag{4}$

Each of the samples $s_j(0)$ represents a complex sample of a beam coming from some direction in the sky. The image, however, consists of a measurement of the power in each beam. The power in each beam may be determined by multiplying each sample in the vector s(0) by its complex conjugate. This calculation is the same as the diagonal of the outer product of s(0) with itself.

$$[s(0)s(0)^*] = diag[[s(0)][s(0)]^H]$$
(5)

Substituting for s(0) from (4):

$$\left[s(0)s(0)^*\right] = diag\left[\left[D\right]\left[v(0)\right]\left[v(0)\right]^H\left[D\right]^H\right]$$
(6)

But the outer product of v(0) with itself is the baseline matrix:

$$\left[s(0)s(0)^*\right] = diag\left[\left[D\right]\left[B(0)\right]\left[D\right]^H\right]$$
(7)

Where H represents the conjugate transpose of the vector or matrix and * represents its complex conjugate. The sum of two such image samples produces the following result.

$$[s(0)s(0)^*] + [s(1)s(1)^*] = diag[[D][B(0)][D]^H] + diag[[D][B(1)][D]^H]$$

$$(8)$$

$$[s(0)s(0)^*] + [s(1)s(1)^*] = diag[[D][B(0) + B(1)][D]^H]$$

$$(9)$$

In a normal observation, a series of baseline calculations (B matrix) are taken and the results are averaged resulting in the covariance matrix R.

$$R = \frac{1}{m} \sum_{i=0}^{m-1} B(i)$$
(10)

$$\frac{1}{m} \sum_{i=0}^{m-1} [s(i)s(i)^*] = diag[[D][R][D]^H]$$
(11)

Where m represents the number of time samples. The image is the diagonal of the transformed covariance matrix.

The complex vector v that generates a baseline array sample B is an eigenvector of B.

$$[B][v] = [v][v]^{H}[v] = [v]p$$
(12)

Where p is the eigenvalue of B. Note that p is the scalar product of v with itself and represents the total power entering the array.

The product of any vector with a baseline array sample matrix will yield its eigenvector or zero. Assume that baseline sample matrix B_0 is formed from the outer product of v_0 with itself. The inner product of v_0 with any vector not orthogonal to it will yield a complex number a_r .

$$a_{x}[v_{0}] = [v_{0}][v_{0}]^{H}[v_{x}] = [B_{0}][v_{x}]$$
(13)

The sum of two baseline sample matrices has, at most, two non-zero eigenvalues.

Assume the sum has eigenvector v_{e} . From equation (13) above

$$b_{0e}[v_0] + b_{1e}[v_1] = [[B_0] + [B_1]][v_e]$$
(14)

 $\therefore v_e$ must be the weighted sum of v_0 and v_1 .

Let
$$[v_e] = a_{0e}[v_0] + a_{1e}[v_1]$$
 (15)

$$a_{0e}a_{00}[v_0] + a_{0e}a_{10}[v_1] + a_{1e}a_{01}[v_0] + a_{1e}a_{11}[v_1] = [[B_0] + [B_1]][a_{0e}[v_0] + a_{1e}[v_1]]$$
(16)

Let k be the eigenvalue of v_e

$$(a_{0e}a_{00} + a_{1e}a_{01})[v_0] + (a_{0e}a_{10} + a_{1e}a_{11})[v_1] = k(a_{0e}[v_0] + a_{1e}[v_1])$$
(17)

$$ka_{0e} = a_{0e}a_{00} + a_{1e}a_{01} \qquad \qquad ka_{1e} = a_{0e}a_{10} + a_{1e}a_{11} \tag{18}$$

$$a_{0e}(k-a_{00}) = a_{1e}a_{01} \qquad a_{1e}(k-a_{11}) = a_{0e}a_{10} \qquad (19)$$

$$a_{0e} = \frac{a_{1e}a_{01}}{k - a_{00}} \qquad \qquad a_{1e} = \frac{a_{0e}a_{10}}{k - a_{11}} \tag{20}$$

$$1 = \frac{a_{01}a_{10}}{(k - a_{00})(k - a_{11})} \qquad a_{01} = a_{10}^*$$
(21)

$$k^{2} - k(a_{11} + a_{00}) + a_{00}a_{11} = a_{01}a_{10}$$
(22)

$$k^{2} - k(a_{11} + a_{00}) + a_{00}a_{11} - a_{01}a_{10} = 0$$
(23)

$$k = \frac{a_{11} + a_{00} \pm \sqrt{(a_{11} - a_{00})^2 + 4a_{01}a_{10}}}{2}$$
(24)

Where k is the eigenvalue.

Note that $\sqrt{a_{01}a_{10}}$ is the norm of the covariance between the two sample sets v_0 and v_1 . If the two sample sets are independent and due to noise then this value will be equal to zero and the two eigenvalues will be equal to a_{00} and a_{11} . If, on the other hand, the two sample sets are completely correlated then k=2a, where $a=a_{00}=a_{11}$ and the other eigenvalue will be zero. It is apparent that any point source or image with coherence across its extent will develop a strong vector root.

Knowing the eigenvalues, it is now possible to determine the corresponding eigenvectors. From (15)

$$\frac{1}{a_{0e}} [v_e] = [v_0] + \frac{a_{1e}}{a_{0e}} [v_1]$$
(25)

From (20)

$$\frac{a_{1e}}{a_{0e}} = \frac{a_{10}}{k - a_{11}} \tag{26}$$

Since an eigenvector times a scaler is still an eigenvector,

$$[v_t] = [v_0] + \frac{a_{10}}{k - a_{11}} [v_1]$$
(27)

 v_t can now be turned into a vector root.

$$k = [v_e]^H [v_e]$$
(28)

$$\begin{bmatrix} v_e \end{bmatrix} = \frac{k \begin{bmatrix} v_t \end{bmatrix}}{\begin{bmatrix} v_t \end{bmatrix}^H \begin{bmatrix} v_t \end{bmatrix}}$$
(29)

An interference identification algorithm: An algorithm suggests itself for collecting baseline data for images with coherence across their extent without computing all of the baselines. It is only necessary to collect the vector square root or eigenvector of the baseline matrix that corresponds to the UV data that exhibits the most coherence from one measurement to the next. This method is very similar to the method proposed by Amir Leshem et al¹ for identifying interference sources. This method does not require the use of the covariance matrix and the identification of its N eigenvectors where N equals the number of antennas.

After the first two samples of array data are gathered calculate all of the inner products of the two sample vectors: a_{00} , a_{11} , and a_{01} . Using these numbers determine the biggest eigenvalue k and its corresponding eigenvector as in 24 and 27. After scaling the eigenvector as in 29 so that its inner product square is equal to its eigenvalue, save it as the vector square root of the baseline matrix corresponding to the interference. Save this interference vector and use it with the next sample vector to repeat the process.

The algorithm, as it has been presented, has a threshold problem. In a low interference to noise environment the norm of the cross covariance between the two sample vectors becomes very small making it difficult to determine which of the vector roots to keep. When the wrong choice is made the image of the interference deteriorates. The threshold becomes lower for bigger arrays since the cross covariance measurement is then based on more samples. Current simulations suggest a threshold of about .01 for an array of 100 elements.

If the astronomical source is a point source then there is a danger of confusion between the astronomical source and the interference. An extended source in radio astronomy does not exhibit coherence across its extent. It behaves like many independent noise sources. Each independent source produces its own signature on the array. In this situation the interference may be more easily identified.

An interference excision algorithm: In order to extract interference a variation on the above algorithm may be used. This algorithm is very similar to the "Spatial Filtering" proposed by Amir Leshem et al¹. Using the norm of the cross correlation to establish a threshold, measure the signature of the interfering source. Repeat the above algorithm to determine the largest eigenvalue and corresponding vector root. Find, as well, the smaller eigenvalue and vector root and pass its vector root onto a standard correlator. The smaller vector root represents signal that is orthogonal to the interference.



figure 1

This figure represents a point spread function of a point source at channel 75 of a 1000 point discrete Fourier transform of a sparse set of 100 randomly distributed points.



figure 2

This figure is the signal of figure 1 added to noise. 50 frames were averaged together. The signal to noise ratio was .01.



figure 3

Each of the 50 frames of the signal of figure 2 was treated with the first algorithm. The coherent signal was successfully removed from the noise background. The signal to noise ratio was .01.



figure 4

This figure shows the successful removal of an interfering coherent signal from noise which could represent an extended astronomical source. The second algorithm was applied to each of the 50 frames of figure 2. The orthogonal eigenroot was passed onto a standard correlator to produce the above result.



figure 5

Random frequencies about the coherent signal of figure 2 were added to represent the effect of an extended source with an interference signal in the middle. As in the other figures, 50 frames were processed.



figure 6

This figure shows the successful isolation of the interfering signal.





The interference removal algorithm passed the extended source and the background noise on to the standard correlator while eliminating the interference.



figure 8

This figure represents the results of an experiment to show that the algorithms distinguish between sources that are coherent across their extent and those that are not. Two point sources are shown which are coherent with each other against a background of noise. 50 frames have been processed.



figure 9

The first algorithm successfully isolates the two sources of figure 8.





The second algorithm removes the coherent signal and passes the noise onto a standard correlator for analysis.

² Simon S. Haykin, "Adaptive Filter Theory" Prentice Hall 1991

¹ Amir Leshem, Alle-Jan van der Veen and Albert-Jan Boonstra "Multichannel Interference Mitigation Techniques in Radio Astronomy" The Astrophysical Journal Supplement Series, 131:355-373, 2000 November