Recovering RFI contributions to visibilities between sky pointing antennas from correlations with and between reference antennas

R.Nityananda, notes of discussions with Ue-Li Pen
NCRA-TIFR, Pune , India, CITA, St.Georges St. Toronto, Canada

July 28, 2010

Abstract

The problem of recovering the RFI contribution to the correlations between sky pointing antennas from their correlations with reference antennas, and the correlations of the reference antennas among themselves, is formulated and a solution based on low rank of the RFI contribution is given.

1 Introduction

Under some simplifying assumptions which have been checked in some cases, each RFI source shows up in visibility matrices as a rank one piece. Considerable improvement in data quality can result from removal of a few dominant RFI sources As proposed and implemented on a trial basis by the GMRT EoR experimenters, this opens up the possibility of using 'reference' antennas - henceforth 'r', whose feeds are pointed towards the astronomical (henceforth 'a') antennas, to estimate the RFI contribution to the visibilities measured among the 'a' antennas. This note is based on my discussions with Ue-Li Pen and puts down the basic formula giving the estimator in terms of the 'rr' (for reference reference) and 'ra' for(astronomical astronomical) correlations.

2 Notation

The total number of antennas is $N = R+A$, and the matrix M of correlations at a given time stamp is given by

$$M = \sum_{\alpha} s_\alpha O_\alpha I_\alpha^\dagger \quad (1)$$

The upper case O 's are r orthogonal vectors in the output space, and the I 's are r orthogonal vectors in the input space, the upper case letters signify that these are the strong (say RFI) modes present, r of them.

In this notation, the extra signal is written as $N = \epsilon a b^\dagger$ with unit vectors a and b and epsilon the perturbation parameter. We decompose a as $\mu e + (1-\mu^2)^{1/2} E$ with E a unit linear combination of the E's and e orthogonal to all the E's, and μ real, nonegative, and less than or equal to one. Likewise, we write $b = \nu i + (1-\nu^2)^{1/2} I$.

3 Statement of the problem

We now consider the matrix $M + N$, i.e RFI plus astronomical source. This had better be of rank $r + 1$ - if it is of rank r, then we can go home because the signal has entirely been swallowed - a new SVD will just show r strong modes. However, in the general case, when the rank is now $r + 1$, we will now get one small singular value, of size $\propto \epsilon$, and it is reasonable to think of that as a filtered version of the signal. The question is then how much of the signal power survives this filtering.

4 Claim and justification

The claim is that the filtered part of the signal will be, with corrections of order ϵ^2, nothing but $\epsilon\mu\nu e i^\dagger$. The analytical rationalisation (note the avoidance of the word "proof") goes as follows. If we do the SVD on $M+N$, we expect corrections to the existing U, S, V of order ϵ, plus the addition of a new singular value and two new unit vectors in the input and output space. We now construct the pertubed SVD (with errors of order ϵ^2 consistent with the claim above. We do this by tackling the four terms, EI, ei, eI, Ei in the expansion of $ab^\dagger$ one by one

The first step is to add $\epsilon(1-\mu^2)^{1/2}(1=\nu^2)^{1/2}EI^\dagger$ to the original matrix and do a new SVD. The matrix remains of rank r, since E and I are linear combinations of the existing E_α and I_α. for brevity we continue to refer to the new U, S, V at this stage by the same symbols, even though they differ from the original ones by terms of order ϵ- The new E, I span the original r-dimensional subspace.

The $(r+1)^{th}$ term in the SVD expansion corresponding to the $(r+1)^{th}$ singular value is chosen to be $\mu\nu e i^\dagger$. This is consistent with the definition of SVD since e and i are orthogonal to the r dimensional space of E's and I's respectively.

The next step in the construction involves changes to U (and V) of order ϵ, which takes them out of their subspace. This is achieved by adding a piece of e of order *epsilon* to each of the r E_α's. this only spoils their orthonormality to order ϵ^2. (It does spoil normality to e to order ϵ, which we will deal with later). Similarly, add a piece of i of order *epsilon* to each of the I's. The coefficients are chosen to deal with the mixed terms like $\epsilon\mu(1-\nu)^{1/2}eI^\dagger$. It is easy to see that we can generate precisely these terms by choosing the coefficients with which we add e to E_α's. For example if I has coefficents c_α when expanded in terms of I_α, then we have to add a term proportional to $c_\alpha e$ to each E_α, which will then combine with I_α in the SVD to give the desired result.

The last small refinement is to rotate e and i by angles of the order of ϵ to restore orthogonality to the new E and i respectively. The price is an error of order ϵ^2 in our goal of representing $ab^\dagger$ by and SVD, which you were warned about anyway.

5 Numerical work

Its easy to fool oneself and some others with arguments of this kind. But octave is too dumb to be fooled, it just does what you tell it to. And I told it to do exactly what is described above - add a rank one piece to a rank r matrix, redo SVD, and pull out the smallest singular value and vectors. And lo and behold, they agreed with $\epsilon\mu\nu$ and e and i!