Effect of SVD RFI filtering on a weak signal

R.Nityananda
NCRA-TIFR, Pune , India, CITA, St.Georges St. Toronto, Canada

July 28, 2010

Abstract

The effect filtering of strong RFI modes detected by SVD on an additional weak -e.g astronomical- signal is assessed analytically and numerically in this note, which is based on ongoing discussions (Sept 2009) with Ue-Li Pen.

1 Introduction

In the GMRT EoR experiment, an idealised RFI source shows up in visibility - time matrices as rank 1 object, being the product of a fixed footprint in baseline-channel-polarisation space - channel space for short - and a fixed time template. An astronomical source has its own, in general different footprint in both these spaces. Its channel footprint varies with time since the delays between antennas change as one tracks, so its contribution to the visibility does not factorise in the above way. (In fact, this is even true for the RFI source because of feed movement but this can be minimised by taking short data stretches). For simplicity we first consider a rank r set of strong RFI sources plus an added weak rank 1 source, In general we then get rank $r+1$ if the original rank was r

2 Notation

A MXN matrix of rank r when expressed in the SVD form can be thought of as a sum of r rank 1 matrices, each the outer product of unit vectors in its output and input spaces, weighted by the singular value. These vectors are orthogonal, and together form the rows of the unitary matrices U and V in the SVD. which is defined by $M = USV^\dagger$. We thus have

$$M = \sum_\alpha s_\alpha O_\alpha I_\alpha^\dagger \tag{1}$$

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 $N = \epsilon ab^\dagger$ with unit vectors a and b and ϵ 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 "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 neglect of terms 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 ϵ 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 ϵ to each of the I's. The coefficients are chosen to deal with the mixed terms like $\epsilon\mu(1-\nu)^{1/2} e I^\dagger$. 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 a programming language like 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 small rank one piece to a rank r matrix, redo SVD, and pull out the smallest singular value and associated vectors. And lo and behold, they agreed with $\epsilon\mu\nu$ and e and i!

6 Comments

This basic result allows one to asess the signal loss caused by RFI removal. In practice, the data matrix is of maximal rank - RFI modes come in all sizes, and hence one needs to decide at what stage one should truncate the RFI model before subtracting. A Wiener like criterion seems reasonable - as one subtracts more and more modes, one loses signal, but one loses noise as well, and the optimum can be found, presumably roughly when the sum of the remaining RFI modes equals the part of the signal expected to survive at that stage.