A linear least squares algorithm for interferometric gain and phase calibration

1 Introduction

In radio interferometry, it is common practice to observe a point source calibrator at the field centre and thereby obtain the complex gains g_i of the antennas. In the ideal situation, the expected visibility is a constant for all baselines. Hence the measured visibilities are related to the gains by

$$V_{ij} = K g_i \; g_j^* + \text{noise} \tag{1}$$

The scale factor K would depend on the known flux of the calibrator. For simplicity, we absorb it into the left hand side in the following.

Equation (1) represents an overdetermined system for N gains and $N-1$ phase differences, and should be treated as a problem of minimising χ^2 since there is noise present. It is usually solved by an iterative method or by nonlinear least squares.

2 The linear least squares formulation and the 2π problem

In this note, we suggest and implement a *linear* least squares algorithm for obtaining complex gains, which has significant advantages of speed and simplicity. The first step is to take logarithms of the visibilities and the gains, i.e to define

$$V_{ij} = \exp\ (\gamma_{ij} + i\epsilon_{ij}); g_i = \exp\ (\beta_i + i\alpha_i) \tag{2}$$

Substituting (2) into (1) and equating the exponents, we get the following set of linear equations.

$$\gamma_{ij} + i\epsilon_{ij} = \beta_i + \beta_j + i(\alpha_i - alpha_j) + \text{noise} \tag{3}$$

We note the follwing regarding (3)

1) The noise on the logarithm of the visibility is not, in general, zero mean gaussian. However, in the limit of small noise, we can write

$$\delta(\gamma + i\epsilon) = \delta V/V = (\delta(Re(V) + iIm(V))/V \tag{4}$$

and one finds that the noise on γ and ϵ is indeed gaussian, but with a variance proportional to $|V|^2$. Thus, one can try and solve (3) by linear least squares with weights proportional to $|V|^2$.

2) The solution for amplitudes, i.e for β_i is straightforward

3) The solution for the phases α_i is *not* straightforward because the logarithm is a multivalued function. There could be an, in general different, integer multiple of 2π on each right hand side in (3). In practical computations, each ϵ_{ij} is obtained using the "atan2" i.e two argument arctangent function, which returns a value of phase between $-\pi$ and π, given the real and imaginary parts of a complex number. A small change in the visibility near the negative real axis can cause a phase jump of 2π in one of the ϵ 's. This cannot, in general, be accomodated by addition of multiples of 2π to the parameters of the solution, viz. α 's. (We have verified this numerically in actual cases, as well as analytically).

Perhaps this problem is responsible for this approach of taking logarithms not being pursued for phase determination. We show in the next section that this problem can be solved and the simplicity and speed of linear least squares regained.

3 Phase unwinding and rewinding

Clearly, the problem mentioned in the previous section would disappear if all the gains g_i had phases lying between $-\pi/2$ and $+\pi/2$, and hence the visibilities would have phases which could never cross π and undergo jumps. This desirable state of affairs can be achieved as follows. Define an auxiliary, fictitious problem whose gains are g' related to the real gains by the equation

$$g_i = g'_i \exp\ (i\zeta_i) \tag{5}$$

Substituting (5) into (1), we get

$$V_{ij} = g'_i g'^*_j \exp\ (i(\zeta_i - \zeta_j))$$

We can move the phase terms to the left side by defining new visibilities $V'_{ij} = V_{ij} \exp\ (-i(\zeta_i - \zeta_j))$. In terms of these new visibilities and new gains, the problem now looks the same as before, viz

$$V'_{ij} = K g'_i\ g'^*_j + \text{noise} \tag{6}$$

Even the noise statistics would be the same as in (1). Now the aim is to choose the parameters ζ_i so that the phases of the new gains g' and hence the new visibilitiesV', as well are all close to zero. There will then be no risk of 2π jumps. To acheive this, it is enough to choose a relatively crude (error less than $\pi/2$)estimate of the phase α_i for the parameter ζ_i, so that the new gain $g'_i = |g_i| \exp\ (i(\alpha_i - \zeta_i))$ have phases close to zero. We call this step "phase unwinding". After this is done, we solve the imaginary part of the system of equations (3) for the phase differences (absolute phases do not enter into the visibility).

It is quite straightforward to estimate the differnces of phases $\alpha_1 - \alpha_2, \alpha_2 - \alpha_3$,etc. from the phases of the visibilities. viz. ϵ_{12}, etc. One can get the same difference, e.g $\alpha_1 - \alpha_2$ from other combinations like $\epsilon_{13} - \epsilon_{23}$. These different estimates of a given phase difference, say $\alpha_1 - \alpha_2$, are "vector averaged" (i.e sin and cos of this quantity are separately averaged and atan2 is applied to the pair). Setting one value, say α_1, to zero as a phase reference, the remaining phases are estimated and sent in as ζ 's for phase unwinding.

4 Removal of amplitude bias at large noise level

The scheme described above was tested against both the iterative and nonlinear least squares algorithms, and on both simulated and real data. It worked well, giving very similar (to within a few per cent) χ^2, except in situations where the noise was large or (equivalently) there were small visibilities due to low gain on some antennas. These are precisely the situations where the Taylor expansion used above in equation (4) to discuss the noise statistics breaks down. One can see that a normal distribution for the error on the parameter γ in (3), which is the logarithm of $|V|$ will actually give a distribution for $|V|$ which is skewed to higher values, and vanishes at $|V| = 0$. This was indeed found in the simulation results which showed systematic bias towards higher values of the amplitude of the antenna complex gain, at large noise levels, with the problem being most serious for "dead" antennas. This problem was greatly mitigated by rescaling all the gains emerging from this algorithm, minimising the χ^2 vs. the measured visibilities with respect to the scale factor. Once this added scaling feature is incorporated, and dead antennas are removed from the solution, the problem of biased amplitudes disappears for all practical purposes.

5 Conclusion

This algorithm, provisionally called "lantsol" with the l standing for logarithm linear, or lazy, ran in about one hundredth the time of a full nonlinear least sqares solution, which was built around code from 'Numerical Recipes (and hence was not tailored to this specific problem). The ability to generate phase solutions for all frequency channels in real time might be useful in some applications involving the study of system performance or ionospheric behaviour.