Weight computation in NDPPP

Sebastiaan van der Tol

December 7, 2011

1 Introduction

This document describes how NDPPP computes the weights when autoweight = True and how the weights are handled in averaging. It also explains briefly why the computation is done that way.

2 Data Model

The observed visibilities $\hat{\mathbf{v}}$ are modeled by

$$\hat{\mathbf{v}} = \mathbf{f}(\boldsymbol{\theta}) + \mathbf{n},$$

where $\mathbf{f}$ is the Measurement Equation, $\boldsymbol{\theta}$ are the parameters and $\mathbf{n}$ is the Gaussian noise. We assume the noise contributions to the visibilities are independent of eachother. That means that the covariance matrix of the noise, defined by

$$\mathbf{R} = \mathrm{E}\left[\mathbf{n}\mathbf{n}^{\mathrm{H}}\right],$$

is a diagonal matrix.

3 Calibration

The Maximum Likelihood estimate of the parameters is given by

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} (\mathbf{f}(\boldsymbol{\theta}) - \hat{\mathbf{v}})^{\mathrm{H}} \mathbf{R}^{-1} (\mathbf{f}(\boldsymbol{\theta}) - \hat{\mathbf{v}})$$

Since $\mathbf{R}$ is diagonal this reduces to

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \sum_{i=0}^{N_v - 1} w_i \| f_i(\boldsymbol{\theta}) - \hat{x}_i \|^2$$

where N_v is the number of visibilities in the calibration data block, and where w_i is the weight given by

$$w_i = \frac{1}{\sigma_i^2}$$

4 Estimating the weights

To set the weights we need to know the variance of the visibilities. The variance of visibility $v_{k,l}$ over baseline $k-l$ is given by

$$\sigma_{kl}^2 = \sigma_k^2 \sigma_l^2 / N_{int},$$

where σ_l^2 and σ_k^2 are the autocorrelations of stations k, l respectively and N_{int} are the number of samples in an integration interval.

The autocorrelations can be estimated from the data. An estimate of the variance is given by

$$\sigma_l^2 \approx v_{ll}$$

Note that the estimate of the variance will be more accurate when based on a longer integration period, at least as long as the noise is stationary over that period. On the other hand, estimation of the weights is not very critical. The information is in the visibilities, the weights only determine how efficiently this information is being used. For practical reasons we have choosen to estimate the weights from the same integration period as the visibilities.

5 Averaging

After the integration by the correlator the data might be averaged further in subsequent processing steps. In this case we do not estimate the noise variance from the data again. The weights are assumed to be exact and the weights of the averaged data are computed from them.

For a set of visibilities with the same expected value the Maximum Likelihood estimate of the expected value is given by

$$\overline{v} = \arg\min_v (\mathbf{v} - \mathbf{1}v)^{\mathrm{H}} \mathbf{R}^{-1} (\mathbf{v} - \mathbf{1}v).$$

Because $\mathbf{R}$ is diagonal this reduces to

$$\overline{v} = \arg\min_v \sum_{i=0}^{N_{int1}-1} w_i \|v_i - v\|^2.$$

The minimum can be found by taking the derivatieve and equating it to zero, leading to

$$\sum_{i=0}^{N_{int1}-1} w_i (v_i - v) = 0$$

Solving the equation above leads to

$$\overline{v} = \frac{\sum_{i=0}^{N_{int1}-1} w_i v_i}{\sum_{i=0}^{N_{int1}-1} w_i}. \tag{1}$$

This shows that when averaging the weights need to be taken into account to obtain the minimum variance average.

The variance of this average is given by

$$\frac{\sum_{i=0}^{N_{int1}-1} w_i^2 \sigma_i^2}{\left(\sum_{i=0}^{N_{int1}-1} w_i\right)^2} = \frac{\sum_{i=0}^{N_{int1}-1} w_i}{\left(\sum_{i=0}^{N_{int1}-1} w_i\right)^2} = \frac{1}{\sum_{i=0}^{N_{int1}-1} w_i}$$

which leads to the following weight

$$w = \sum_{i=0}^{N_{int1}-1} w_i \tag{2}$$

The weight for the averaged data is the sum of the weights of the components