ngEHT Model Fitting Precision Estimates

Avery Broderick

June 2022

1 Introduction

Typically, model parameters can be constrained with far greater precision than implied by the nominal beam in VLBI experiments. Here we describe the formalism underlying the FisherForecast package used to estimate the precision with which model parameters can be recovered from VLBI data.

We will leverage the fact that we know the "truth" model, assume Gaussian likelihoods and Gaussian posteriors, consider only fits to complex visibilities, and assume that the complex gains can be recovered.

2 Joint Uncertainty Estimates

We make use of the standard Fisher matrix approach, Taylor expanding the log-likelihood about the "truth", i.e.,

$$\mathcal{L}(p+\delta p) = \left.\mathcal{L}\right|_p + \left.\frac{\partial \mathcal{L}}{\partial p}\right|_p \cdot \delta p + \frac{1}{2}\delta p \cdot \left.\frac{\partial^2 \mathcal{L}}{\partial p \partial p}\right|_p \cdot \delta p + \dots \tag{1}$$

where

$$\mathcal{L}(p) = -\sum_j \frac{|V_j - W_j|^2}{2\sigma_j^2}, \tag{2}$$

in which W_j is the model visibility and depends on the parameters p. The second and third terms are

$$\left.\frac{\partial \mathcal{L}}{\partial p}\right|_p = \sum_j \left[\frac{\partial W_j^*}{\partial p} \frac{(V_j - W_j)}{2\sigma_j^2} + \frac{(V_j - W_j)^*}{2\sigma_j^2} \frac{\partial W_j}{\partial p} \right], \tag{3}$$

and

$$\left.\frac{\partial^2 \mathcal{L}}{\partial p \partial p}\right|_p = \sum_j \left[\frac{\partial^2 W_j^*}{\partial p \partial p} \frac{(V_j - W_j)}{2\sigma_j^2} + \frac{(V_j - W_j)^*}{2\sigma_j^2} \frac{\partial^2 W_j}{\partial p \partial p} - \frac{1}{\sigma_j^2} \frac{\partial W_j^*}{\partial p} \frac{\partial W_j}{\partial p} \right]. \tag{4}$$

Averaging over thermal realizations, all of the $V_j - W_j$ terms vanish, giving

$$\left.\frac{\partial \mathcal{L}}{\partial p}\right|_p = 0 \quad \text{and} \quad \left.\frac{\partial^2 \mathcal{L}}{\partial p \partial p}\right|_p = -\sum_j \frac{1}{\sigma_j^2} \frac{\partial W_j^*}{\partial p} \frac{\partial W_j}{\partial p} \equiv -M. \tag{5}$$

This latter is simply the Fisher matrix, and is convenient because it depends only on the model and the measurement response (i.e., the coordinates at which data is measured and σ_j).

Therefore, the "typical" posterior, set by the mean $\mathcal{L}$, is approximately Gaussian and given by,

$$P(p) \approx N \exp\left(-\frac{1}{2}\delta p \cdot M \cdot \delta p\right), \tag{6}$$

where N contains the relevant normalization terms (which include $\exp(\mathcal{L}|_p)$. Once a model is selected and the observations specified, M can be constructed and the Posterior immediately estimated. The M is simply the inverse of the standard covariance.

3 Ideal and Marginalized Estimates

Two classes of uncertainties are relevant:

Ideal estimates associated with fixing all other model parameters. By asserting strong priors on all other model aspects, this provides an assessment of the optimal situation: what is the maximum precision attainable even in principle.

Marginalized estimates after integrating the posterior over all of the other model parameters. These provide a typical assessment where the priors are uninformative.

We may generate these analytically immediately once M is defined. We collect expressions for when estimates for the precision of a single parameter is sought (single) and the joint precision of two parameters is sought (joint).

3.1 Single Parameter Estimates

We begin with the case where we seek only the precision possible on the ith parameter, p_i. The ideal uncertainty estimate is

$$\sigma_{p_i} = M_{ii}^{-1/2}. \tag{7}$$

Marginalizing over the other parameters requires performing a Gaussian integral, which results in

$$\sigma_{p_i}^m = \left[M_{ii} - \mu_i^T m_i^{-1} \mu_i\right]^{-1/2}, \tag{8}$$

where μ_i is the $(n-1)$-dimensional vector constructed from the ith column of M excluding the ii position, and m_i is the $(n-1)$-dimensional square matrix constructed by excluding the ith row and column of M. This expression may be trivially derived by "completing the square" in the exponent.

3.2 Joint Parameter Estimates

Frequently, parameters will be correlated, and thus joint estimates are desirable. We restrict ourselves to two-dimensional parameter estimates for which we will require an appropriate analog of M. That is, we wish the joint probability distribution of p_i and p_j, which we write as

$$P(p_i, p_j) = \exp\left[-\frac{1}{2}\left(N_i \delta p_i^2 + N_j \delta p_j^2 + 2C_{ij}\delta p_i \delta p_j\right)\right] \tag{9}$$

In the ideal case,

$$N_i = M_{ii}, \quad N_j = M_{jj}, \quad \text{and} \quad C_{ij} = M_{ij} = M_{ji}. \tag{10}$$

In the marginalized case,

$$\begin{aligned} N_i^m = M_{ii} - \bar{\mu}_{i,ij}^T \bar{m}_{ij}^{-1} \bar{\mu}_{i,ij}, \quad N_j^m = M_{jj} - \bar{\mu}_{j,ij}^T \bar{m}_{ij}^{-1} \bar{\mu}_{j,ij}, \\ \text{and} \quad C_{ij}^m = M_{ij} - \frac{1}{2}\bar{\mu}_{i,ij}^T \bar{m}_{ij}^{-1} \bar{\mu}_{j,ij} - \frac{1}{2}\bar{\mu}_{j,ij}^T \bar{m}_{ij}^{-1} \bar{\mu}_{i,ij}, \end{aligned} \tag{11}$$

where $\bar{\mu}_{k,ij}$ is the $(n-2)$-dimensional vector constructed from kth column of M excluding the ith and jth rows, and $\bar{m}_{ij}$ is the $(n-2)$-dimensional square matrix constructed from M by excluding the ith and jth rows and columns. Again, this may be readily derived by completing the square.

4 FisherForecast Package

The FisherForecast package is simply an implementation of the above formalism and a handful of models. These currently include a simple ring model, a "themage", a symmetric Gaussian, and sums thereof. The package is contained in `fisher_package.py` and examples for a ring+themage model and binary Gaussians are in `ring_example.py` and `binary_example.py`, respectively. Current dependencies include Matplotlib, Numpy, SciPy, ehtim (to read in observation files), and ThemisPy (to enable support for all ThemisPy `model_image` class objects).