Error Computation for Emulation

April 14, 2020

1   Introduction

Our goal is to compute the likelihood function used to estimate model parameters with our emulator. The log-likelihood $\mathcal{L}$ can be defined as

$$\mathcal{L} = -\frac{1}{2}\left(\frac{y_{\mathrm{pred}} - y_{\mathrm{obs}}}{y_{\mathrm{obs}}}\right)C^{-1}\left(\frac{y_{\mathrm{pred}} - y_{\mathrm{obs}}}{y_{\mathrm{obs}}}\right) \tag{1}$$

where C is the covariance matrix, and y is the clustering statistic we are emulating. We use y_{obs} to mean the statistic computed on a mock or data. We use y_{pred} to mean the emulator prediction for that statistic. We take y to have P dimensions (radial bins in our case), and C to be a $P \times P$ matrix.

   This document outlines how we compute this covariance matrix C, as well as other error measurements used in our emulation procedure.

2   Notation

The test set contains mock catalogs with $C = 7$ different cosmologies c. There are $B = 5$ boxes (or realizations) b for each cosmology. These are each populated with $H = 100$ different halo occupation distribution (HOD) models h, the same 100 for each cosmology. (There are also 10 random seeds used per HOD, but for now we only use one. We'll worry about this later.)

   All of the covariance matrices in this document are fractional errors. This is consistent with our definition of the likelihood function, which uses the fractional error between the observation and the emulator prediction. We note that generally, it is not obvious whether to divide the physical error by the observation or the prediction to obtain the fractional error. In this case,

the prediction is trained on many observations, so the observation y_{obs} is the more well-measured quantity and we choose it for the denominator.

The statistic is computed on all $7 \times 5 \times 100 = 3500$ mocks. The statistic is averaged over the 5 boxes, so our final test set contains $N = 700$ statistics. We will denote these with index n.

3 Test set error

The error on the test set is computed as follows. For each cosmology, compute the mean statistic over the boxes. Compute the deviation from this mean for each of the boxes, defined as the fractional error between the box statistic and the mean statistic for that cosmology. Do this for every HOD. The test set error is the standard deviation of all of these deviations from the mean.

Breaking this down: For each cosmology and HOD, compute the mean statistic $\bar{y}_{c,h}$ of the B boxes.

$$\bar{y}_{c,h} = \frac{1}{B} \sum_b^B y_{b,c,h} \tag{2}$$

For each cosmology and HOD, compute the deviations from the mean for each box b.

$$d_{b,c,h} = \frac{y_{b,c,h} - \bar{y}_{c,h}}{\bar{y}_{c,h}} \tag{3}$$

Compute the covariance of all of the deviations form the mean.

$$C^{\text{test}} = \frac{1}{BCH - 1} \sum_b^B \sum_c^C \sum_h^H d_{b,c,h} \cdot d_{b,c,h}^{\text{T}} \tag{4}$$

C^{test} is a matrix of size $P \times P$. The square root of the diagonal of this matrix (a.k.a. the standard deviation) is used as an input to the Gassian Process to represent the error on the *training set*. We do this because have a better handle on this error, compared to using the training set, due to the multiple realizations of the test set. We note that the procedure outlined here estimates the sample variance contribution to the error. The random seeds we have not yet included estimate the shot noise contribution; we will update with this later.

4 Emulator performance and error

The overall emulator performance C^{perf} is defined as

$$C^{\text{perf}} = C^{\text{emu}} + C^{\text{test}} \tag{5}$$

We obtain this with the covariance of the fractional error of the test set emulation:

$$f_n = \frac{y_{n,pred} - y_{n,obs}}{y_{n,obs}} \tag{6}$$

$$C^{\text{perf}} = \frac{1}{N-1} \sum_n^N f_n \cdot f_n^{\mathrm{T}} \tag{7}$$

We care about the emulator error σ_{emu} separately, so we use

$$C^{\text{emu}} = C^{\text{perf}} - C^{\text{test}} \tag{8}$$

Both C^{emu} and C^{perf} are matrices of size $P \times P$.

5 Measurement covariance

The covariance between the bins of the statistic is obtained with the Minerva mock catalogs. We compute the statistic on $M = 100$ Minerva mocks, which we index by m. The Minerva covariance matrix C^{minerva} is a $P \times P$ matrix, given by

$$C^{\text{minerva}} = \frac{1}{N-1} \sum_m^M (y_m - \bar{y})(y_m - \bar{y})^T \tag{9}$$

We combine these errors to obtain a final covariance matrix C for our likelihood. The final covariance matrix is

$$C = C^{\text{minerva}} + C^{\text{emu}} \tag{10}$$