Approximating the end of nested sampling

Zixiao Hu, Artyom Baryshnikov, Will Handley

26 September 2023

ABSTRACT

Key words: methods: data analysis – methods: statistical

```
Predicted endpoint: 25054 +/- 242
Progress: [================>########] 72%

_____
lives      |   500 |
phantoms   | 24310 |
posteriors | 18018 |
equals     |   245 |
------------------
ncluster    =  1/1
ndead       =  18018
nposterior  =  18018
nequals     =  249
nlike       =  4159049
<nlike>     =  491.04 (9.82 per slice)
log(Z)      =  -12.55 +/- 0.27
```

Figure 1. Output from PolyChord for a typical nested sampling run. The predicted endpoint, shown in red, is calculated using the method described in this paper.

1 INTRODUCTION

Nested sampling is a multi-purpose algorithm invented by John Skilling which simultaneously functions as a probabilistic sampler, integrator and optimiser (Skilling 2006). It was immediately adopted for cosmology, and is now used in a wide range of physical sciences including particle physics, materials science (Ashton et al. 2022) and machine learning (Higson et al. 2018). The core algorithm is unique for estimating volumes by *counting*, which makes high-dimensional integration feasible. It also avoids problems faced by traditional Bayesian algorithms, such as multi-modality.

The order of magnitude runtime of an algorithm, that is, whether termination is hours or weeks and months away, is of high importance to the end user. Currently, existing implementations of nested sampling either do not give any indication of the runtime, or only provide crude estimates. In the latter case, these are always underpredictions which usually fail to recover the correct order of magnitude.

This paper sets out a more principled manner of endpoint estimation for nested sampling at each intermediate stage, the key idea being to use the existing samples to predict the likelihood in the region we have yet to sample from. Outline of paper.

2 BACKGROUND

2.1 The progress of a nested sampling run

Present some theoretical results on anatomy of nested sampling run.

2.2 Current prior volume

2.3 Current log-likelihood

2.3.1 Dimensionality

Define Bayesian model dimensionality Goal to observe how the number of constrained parameters changes throughout the run Only works at the end of the run, because BMD is a property of the posterior - heat capacity analogy Need to transform intermediate set of samples into a posterior. Useful tool is the transformation using beta. Which beta to use? Several existing temperatures in the literature - microcanonical temperature, canonical temperature used by Habeck. Propose a temperature from Bayesian inference.

Ridge geometries contain phase transitions in nested sampling, because the dimensionality of the samples increases throughout the run.

As an example, consider an elongated Gaussian in a unit hypercube.

3 ENDPOINT PREDICTION

Discussion of how to get endpoint and the necessity of extrapolation. Compare to UltraNest/dynesty slab methods and why they underpredict.

The key observation is that the Bayesian model dimensionality is the equivalent dimension of the posterior if it were actually Gaussian. Fitting a Gaussian of this dimension to the likelihood profile should be a reasonable approximation to the true distribution.

Show fits Find BMD Fit data to find sigma Extrapolate using a Gaussian likelihood with this parameterisation

REFERENCES

Ashton G., et al., 2022, Nature Reviews Methods Primers, 2
Higson E., Handley W., Hobson M., Lasenby A., 2018, Monthly Notices of the Royal Astronomical Society
Skilling J., 2006, Bayesian Analysis, 1