anesthetic: nested sampling visualisation

Will Handley[1, 2, 3]

1 Astrophysics Group, Cavendish Laboratory, J.J.Thomson Avenue, Cambridge, CB3 0HE, UK **2** Kavli Institute for Cosmology, Madingley Road, Cambridge, CB3 0HA, UK **3** Gonville & Caius College, Trinity Street, Cambridge, CB2 1TA, UK

Summary

`anesthetic` is a Python package for processing nested sampling runs, and will be useful for any scientist or statistician who uses nested sampling software. `anesthetic` unifies many existing tools and techniques in an extensible framework that is intuitive for users familiar with the standard Python packages, namely NumPy, SciPy, Matplotlib and pandas. It has been extensively used in recent cosmological papers (W. Handley & Lemos, 2019a, 2019b).

Nested sampling

Nested sampling (Skilling, 2006) is an alternative to Markov-Chain-Monte-Carlo techniques (Hastings, 1970). Given some data D, for a scientific model M with free parameters θ, Bayes theorem states:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}.$$

Traditional MCMC approaches ignore the Bayesian evidence $P(D)$ and instead focus on the problem of generating samples from the posterior $P(\theta|D)$ using knowledge of the prior $P(\theta)$ and likelihood $P(D|\theta)$. Nested sampling reverses this priority, and instead computes the evidence $P(D)$ (the critical quantity in Bayesian model comparison (Trotta, 2008)), producing posterior samples as a by-product. Nested sampling does this by evolving a set of live points drawn from the prior under a hard likelihood constraint which steadily increases, causing the live points to contract around the peak(s) of the likelihood. The history of the live-point evolution can be used to reconstruct both the evidence and posterior samples, as well as the density of states and consequently the full partition function.

Current publicly available implementations of nested sampling include MultiNest (Feroz, Hobson, & Bridges, 2009), PolyChord (W. J. Handley, Hobson, & Lasenby, 2015a, 2015b; Higson, 2018), DNest (Brewer & Foreman-Mackey, 2018) and dynesty (Speagle, 2019), all of which have been incorporated into a wide range of cosmological (Brinckmann & Lesgourgues, 2019; Antony Lewis & Bridle, 2002; Zuntz et al., 2015) and particle physics (The GAMBIT Scanner Workgroup:, 2017) codes.
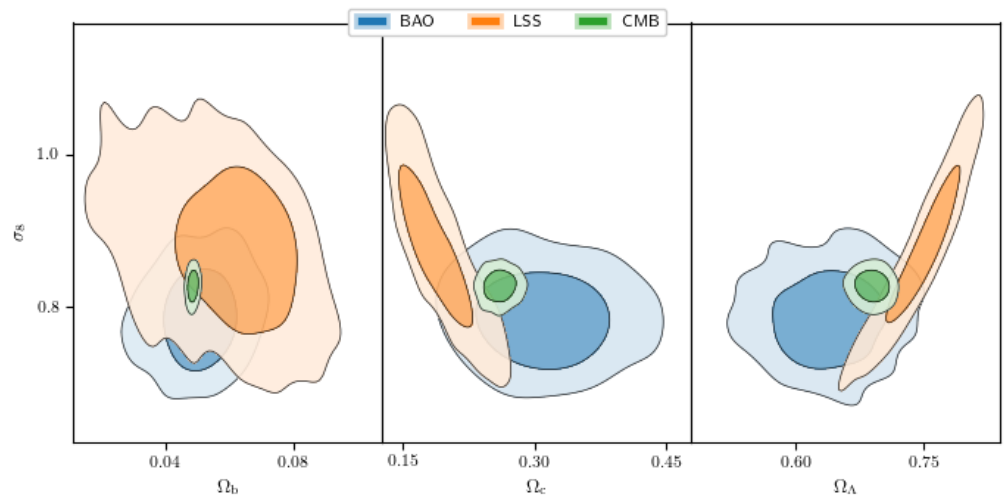
Figure 1: Marginalised posterior plots produced by `anesthetic`. The x axes indicate the fraction of normal matter, dark matter and dark energy respectively, whilst the y-axis is the amplitude of mass fluctuation in our late-time universe. The three measurements were performed using measurements of baryonic acoustic oscillations, large scale structure and the cosmic microwave background (W. Handley & Lemos, 2019a). It is an open cosmological and statistical questions whether the LSS and CMB are consistent with one another.

aNESThetic

`anesthetic` acts on outputs of nested sampling software packages. It can:

1. Compute inferences of the Bayesian evidence (Trotta, 2008), the Kullback-Leibler divergence (Kullback & Leibler, 1951) of the distribution, the Bayesian model dimensionality (W. Handley & Lemos, 2019b) and the full partition function.
2. Dynamically replay nested sampling runs.
3. Produce one- and two-dimensional marginalised posterior plots (Figure 1).

A subset of computations from item 1 is provided by many of the nested sampling software packages. `anesthetic` allows you to compute these independently and more accurately, providing a unified set of outputs and separating these computations from the generation of nested samples.

Item 2 is useful for users that have experienced the phenomenon of 'live point watching' – the process of continually examining the evolution of the live points as the run progresses in an attempt to diagnose problems in likelihood and/or sampling implementations. The GUI provided allows users to fully reconstruct the run at any iteration, and examine the effect of dynamically adjusting the thermodynamic temperature.

Finally, it is important to recognise that the functionality from item 3 is also provided by many other high-quality software packages, such as getdist (Anthony Lewis, 2015), corner (Foreman-Mackey, 2016), pygtc (Bocquet & Carter, 2016), dynesty (Speagle, 2019) and MontePython (Brinckmann & Lesgourgues, 2019). `anesthetic` adds to this functionality by:

- Performing kernel density estimation using the state-of-the-art fastkde (T. A. O'Brien, Kashinath, Cavanaugh, Collins, & O'Brien, 2016) algorithm.

- Storing samples and plotting grids as a weighted `pandas.DataFrame`, which is more consistent with the scientific Python canon, allows for unambiguous access to samples and plots via their reference names, and easy definition of new parameters.
- Using a contour colour scheme that is better suited to plotting distributions with uniform probability, which is important if one wishes to plot priors along with posteriors.

The source code for `anesthetic` is available on GitHub, with its automatically generated documentation at ReadTheDocs and a pip-installable package on PyPi. An example interactive Jupyter notebook is given using Binder (Jupyter et al., 2018). Continuous integration is implemented with Travis and Circle.

Acknowledgements

References

Bocquet, S., & Carter, F. W. (2016). Pygtc: Beautiful parameter covariance plots (aka. Giant triangle confusograms). *The Journal of Open Source Software*, *1*(6). doi:10.21105/joss.00046

Brewer, B., & Foreman-Mackey, D. (2018). DNest4: Diffusive nested sampling in c++ and python. *Journal of Statistical Software, Articles*, *86*(7), 1–33. doi:10.18637/jss.v086.i07

Brinckmann, T., & Lesgourgues, J. (2019). MontePython 3: Boosted MCMC sampler and other features. *Physics of the Dark Universe*, *24*, 100260. doi:10.1016/j.dark.2018.100260

Feroz, F., Hobson, M. P., & Bridges, M. (2009). MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Monthly Notices of the Royal Astronomical Society*, *398*, 1601–1614. doi:10.1111/j.1365-2966.2009.14548.x

Foreman-Mackey, D. (2016). Corner.py: Scatterplot matrices in python. *The Journal of Open Source Software*, *24*. doi:10.21105/joss.00024

Handley, W. J., Hobson, M. P., & Lasenby, A. N. (2015a). POLYCHORD: nested sampling for cosmology. *Monthly Notices of the Royal Astronomical Society*, *450*, L61–L65. doi:10.1093/mnrasl/slv047

Handley, W. J., Hobson, M. P., & Lasenby, A. N. (2015b). POLYCHORD: next-generation nested sampling. *Monthly Notices of the Royal Astronomical Society*, *453*, 4384–4398. doi:10.1093/mnras/stv1911

Handley, W., & Lemos, P. (2019a). Quantifying tension: interpreting the DES evidence ratio. *arXiv e-prints*, arXiv:1902.04029.

Handley, W., & Lemos, P. (2019b). Quantifying dimensionality: Bayesian cosmological model complexities. *arXiv e-prints*, arXiv:1903.06682.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*(1), 97–109. doi:10.2307/2334940

Higson, E. (2018). dyPolyChord: Dynamic nested sampling with PolyChord. *Journal of Open Source Software*, *3*(29), 916. doi:10.21105/joss.00965

Jupyter et al. (2018). Binder 2.0 - Reproducible, Interactive, Sharable Environments for Science at Scale. In Proceedings of the 17th python in science conference. doi:10.25080/Majora-4af1f417-011

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *Ann. Math. Statist.*, *22*(1), 79–86. doi:10.1214/aoms/1177729694

Lewis, A. (2015). Getdist github repository. Retrieved from https://github.com/cmbant/getdist

Lewis, A., & Bridle, S. (2002). Cosmological parameters from CMB and other data: A Monte Carlo approach. *Phys. Rev.*, *D66*, 103511. doi:10.1103/PhysRevD.66.103511

O'Brien, T. A., Kashinath, K., Cavanaugh, N. R., Collins, W. D., & O'Brien, J. P. (2016). A fast and objective multidimensional kernel density estimation method: FastKDE. *Computational Statistics & Data Analysis*, *101*, 148–160. doi:10.1016/j.csda.2016.02.014

Skilling, J. (2006). Nested sampling for general bayesian computation. *Bayesian Analysis.*, *1*(4), 833–859. doi:10.1214/06-BA127

Speagle, J. S. (2019). dynesty: A Dynamic Nested Sampling Package for Estimating Bayesian Posteriors and Evidences. *arXiv e-prints*, arXiv:1904.02180.

The GAMBIT Scanner Workgroup:. (2017). Comparison of statistical sampling methods with scannerbit, the gambit scanning module. *The European Physical Journal C*, *77*(11), 761. doi:10.1140/epjc/s10052-017-5274-y

Trotta, R. (2008). Bayes in the sky: Bayesian inference and model selection in cosmology. *Contemporary Physics*, *49*, 71–104. doi:10.1080/00107510802066753

Zuntz, J., Paterno, M., Jennings, E., Rudd, D., Manzotti, A., Dodelson, S., Bridle, S., et al. (2015). CosmoSIS: Modular cosmological parameter estimation. *Astronomy and Computing*, *12*, 45–59. doi:10.1016/j.ascom.2015.05.005