The Journal of Open Source Software

maxsmooth: Derivative Constrained Function Fitting

Harry T. J. Bevins[1]

1 Astrophysics Group, Cavendish Laboratory, J.J.Thomson Avenue, Cambridge, CB3 0HE, UK

DOI: 10.21105/joss.02540

Software

- Review
- Repository
- Archive

Editor: Pending Editor

Submitted: 31 July 2020
Published: 01 August 2020

Summary

`maxsmooth` is an optimisation routine for fitting derivative constrained functions (DCFs) to data sets. DCFs are a family of functions which have derivatives that do not cross zero in the band of interest. Consequently, they can produce perfectly smooth fits to data sets and reveal non-smooth signals of interest in the residuals. `maxsmooth` utilises the CVXOPT implementation of quadratic programming to perform the constrained fitting rapidly and efficiently.

Statement of Need

The development of `maxsmooth` has largely been motivated by the problem of foreground modelling in Global 21-cm experiments. (Bowman, Rogers, Monsalve, Mozdzen, & Mahesh, 2018; de Lera Acedo, 2019; Price et al., 2018; Singh et al., 2018). Global 21-cm cosmology is the study of the spin temperature of hydrogen gas and its relative magnitude when compared to the Cosmic Microwave Background (CMB) during the Cosmic Dawn and Epoch of Reionisation. During this period in the Universe's history the first stars began to form and the properties of the gas in the Universe changed as it interacted with radiation from the first luminous sources (Barkana, 2016; Furlanetto, Oh, & Briggs, 2006; Pritchard & Loeb, 2012).

The goal of Global 21-cm experiments is to detect this structure in the sky averaged radio spectrum between approximately 50 and 200 MHz. However, the signal of interest is expected to be of order ≤ 250 mK and masked by foregrounds $10^4 - 10^5$ orders of magnitude brighter (Cohen, Fialkov, Barkana, & Lotem, 2017; Cohen, Fialkov, Barkana, & Monsalve, 2019).

Modelling and removal of the high magnitude foregrounds is essential for detection of the Global 21-cm signal. The foregrounds are predominantly composed of smooth spectrum synchrotron and free-free emission from the Galaxy and extragalactic radio sources. Consequently, DCFs provide a powerful alternative to standard polynomials for accurately modelling the foregrounds in such experiments.

While `maxsmooth` has been designed with Global 21-cm cosmology in mind it is equally applicable to any experiment in which the signal of interest has to be separated from comparatively high magnitude smooth signals or foregrounds.

`maxsmooth`

DCFs can be fitted with routines such as Basin-hopping (Wales & Doye, 1997) and Nelder-Mead (Nelder & Mead, 1965) and this has been the practice when fitting Maximally Smooth Functions (MSFs), a specific form of DCF, for 21-cm Cosmology (Sathyanarayana Rao, Subrahmanyan, Udaya Shankar, & Chluba, 2017; Singh & Subrahmanyan, 2019). However, maxsmooth employs quadratic programming to rapidly and efficiently fit DCFs which are constrained such that

$$\frac{d^m y}{dx^m} \geq 0 \text{ or } \frac{d^m y}{dx^m} \leq 0,$$

where for MSFs m, the derivative order, is ≥ 2. An example DCF from the library of 7 built-in to maxsmooth is given by

$$y = \sum_{k=0}^{N} a_k \, x^k,$$

where x and y are the independent and dependent variables respectively and N is the order of the fit with powers from $0 - (N-1)$. The library is not intended to be complete and will be extended by future contributions from users. We find that the use of quadratic programming makes maxsmooth approximately two orders of magnitude quicker than a Basin-hopping/Nelder-Mead approach. We find that maxsmooth can fit higher order fits to a higher degree of quality, lower χ^2, without modification of algorithm parameters as is needed when using Basin-hopping.

maxsmooth rephrases the above condition such that

$$\pm_m \frac{d^m y}{dx^m} \leq 0,$$

where the $\pm$ applies to a given m. This produces a set of discrete sign spaces with different combinations of constrained positive and negative derivatives. In each sign space the associated minimum is found using quadratic programming and then maxsmooth identifies the optimum combination of signs, $\mathbf{s}$ on the derivatives. To summarise the minimisation problem we have

$$\min_{a,\ s} \ \frac{1}{2} \, \mathbf{a}^T \, \mathbf{Q} \, \mathbf{a} \, + \, \mathbf{q}^T \, \mathbf{a},$$
$$\text{s.t.} \ \ \mathbf{G(s)} \, \mathbf{a} \leq \mathbf{0},$$

where we are minimising χ^2, $\mathbf{G(s)a}$ is a stacked matrix of derivative evaluations and $\mathbf{a}$ is the matrix of parameters we are attempting to optimise for. $\mathbf{Q}$ and $\mathbf{q}$ are given by

$$\mathbf{Q} = \mathbf{\Phi}^T \, \mathbf{\Phi} \text{ and } \mathbf{q}^T = -\mathbf{y}^T \, \mathbf{\Phi},$$

here $\mathbf{\Phi}$ is a matrix of basis function evaluations and $\mathbf{y}$ is a column matrix of the dependent data points we are fitting.

The discrete spaces can be searched in their entirety quickly and efficiently or alternatively a sign navigating algorithm can be invoked using maxsmooth which reduces the runtime of the algorithm. Division of the parameter space into discrete sign spaces allows for a more complete exploration of the parameter space when compared to Basin-hopping/Nelder-Mead based algorithms.

The sign navigating approach uses a cascading algorithm to identify a candidate set of signs and parameters for the global minimum. The cascading algorithm starts with a randomly generated set of signs. Each individual sign is then flipped, from the lowest order derivative first, until the objective function decreases in value. The signs associated with the lower χ^2 value then become the optimum set and the processes is repeated until the objective function stops decreasing in value. This is then followed by a limited exploration of the neighbouring sign spaces to identify the true global minimum.

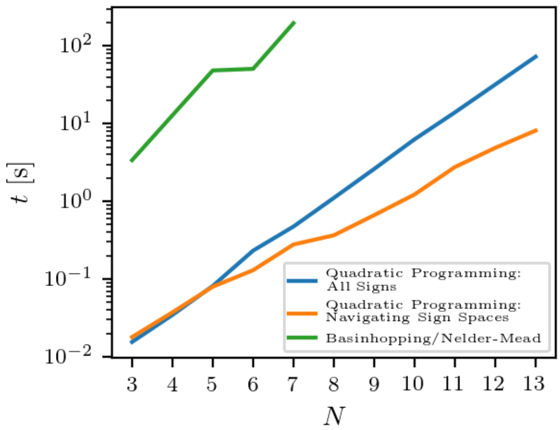

Figure 1: The time taken to fit a polynomial data set following an approximate x^a power law using both `maxsmooth` quadratic programming methods and for comparison a method based in Basin-hopping and Nelder-Mead routines. We show the results using the later method up to $N = 7$ after which the method begins to fail without adjustments to the routine parameters. For $N = 3 - 7$ we find a maximum difference of 0.04% between the optimum `maxsmooth` χ^2 values and the Basin-hopping results. Figure taken from Bevins et al. (2020).

Documentation for `maxsmooth` is available at ReadTheDocs and the code can be found on Github. Version 1.1.0 is available as a PyPI package. Continuous integration is performed with Travis and the associated code coverage can be found at CodeCov.

Acknowledgements

Discussions on the applications of the software were provided by Eloy de Lera Acedo, Will Handley and Anastasia Fialkov. The author is supported by the Science and Technology Facilities Council (STFC) via grant number ST/T505997/1.

References

Barkana, R. (2016). The rise of the first stars: Supersonic streaming, radiative feedback, and 21-cm cosmology. *Physics Reports*, *645*, 1–59. doi:10.1016/j.physrep.2016.06.006

Bevins, H. T. J., Handley, W. J., Fialkov, A., Lera de Acedo, E., Greenhill, L. J., & Price, D. C. (2020). maxsmooth: Rapid maximally smooth function fitting with applications in Global 21-cm cosmology. *arXiv e-prints*, arXiv:2007.14970. Retrieved from http://arxiv.org/abs/2007.14970

Bowman, J. D., Rogers, A. E. E., Monsalve, R. A., Mozdzen, T. J., & Mahesh, N. (2018). An absorption profile centered at 78 megahertz in the sky-averaged spectrum. *Nature*, *555*(7694), 67–70. doi:10.1038/nature25792

Cohen, A., Fialkov, A., Barkana, R., & Lotem, M. (2017). Charting the parameter space of the global 21-cm signal. *Monthly Notices of the Royal Astronomical Society*, *472*(2), 1915–1931. doi:10.1093/mnras/stx2065

Cohen, A., Fialkov, A., Barkana, R., & Monsalve, R. (2019). Emulating the Global 21-cm Signal from Cosmic Dawn and Reionization. *arXiv e-prints*, arXiv:1910.06274. Retrieved from http://arxiv.org/abs/1910.06274

de Lera Acedo, E. (2019). REACH: Radio experiment for the analysis of cosmic hydrogen. *2019 International Conference on Electromagnetics in Advanced Applications (ICEAA)*, 0626–0629. doi:10.1109/ICEAA.2019.8879199

Furlanetto, S. R., Oh, S. P., & Briggs, F. H. (2006). Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe. *Physics Reports*, *433*(4-6), 181–301. doi:10.1016/j.physrep.2006.08.002

Nelder, J. A., & Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer Journal*, *7*(4), 308–313. doi:10.1093/comjnl/7.4.308

Price, D. C., Greenhill, L. J., Fialkov, A., Bernardi, G., Garsden, H., Barsdell, B. R., Kocz, J., et al. (2018). Design and characterization of the Large-aperture Experiment to Detect the Dark Age (LEDA) radiometer systems. *Monthly Notices of the Royal Astronomical Society*, *478*(3), 4193–4213. doi:10.1093/mnras/sty1244

Pritchard, J. R., & Loeb, A. (2012). 21 cm cosmology in the 21st century. *Reports on Progress in Physics*, *75*(8), 086901. doi:10.1088/0034-4885/75/8/086901

Sathyanarayana Rao, M., Subrahmanyan, R., Udaya Shankar, N., & Chluba, J. (2017). Modeling the radio foreground for detection of cmb spectral distortions from the cosmic dawn and the epoch of reionization. *The Astrophysical Journal*, *840*(1), 33. doi:10.3847/1538-4357/aa69bd

Singh, S., & Subrahmanyan, R. (2019). The Redshifted 21 cm Signal in the EDGES Low-band Spectrum. *The Astrophysical Journal*, *880*(1), 26. doi:10.3847/1538-4357/ab2879

Singh, S., Subrahmanyan, R., Udaya Shankar, N., Sathyanarayana Rao, M., Girish, B. S., Raghunathan, A., Somashekar, R., et al. (2018). SARAS 2: a spectral radiometer for probing cosmic dawn and the epoch of reionization through detection of the global 21-cm signal. *Experimental Astronomy*, *45*(2), 269–314. doi:10.1007/s10686-018-9584-3

Wales, D. J., & Doye, J. P. K. (1997). Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. *The Journal of Physical Chemistry A*, *101*(28), 5111–5116. doi:10.1021/jp970984n