The Journal of Open Source Software

Verde: Processing and gridding spatial data using Green's functions

Leonardo Uieda[1]

1 Department of Geology and Geophysics, SOEST, University of Hawai'i at Mānoa, Honolulu, Hawaii, USA

DOI: 10.21105/joss.00944

Software

- Review
- Repository
- Archive

Submitted: 14 September 2018
Published: 14 September 2018

Summary

Measurements made on the surface of the Earth are often sparse and unevenly distributed. For example, GPS displacement measurements are limited by the availability of ground stations and airborne geophysical measurements are highly sampled along flight lines but there is often a large gap between lines. Many data processing methods require data distributed on a uniform regular grid, particularly methods involving the Fourier transform or the computation of directional derivatives. Hence, the interpolation of sparse measurements onto a regular grid (known as *gridding*) is a prominent problem in the Earth Sciences.

Popular gridding methods include kriging, minimum curvature with tension (W. Smith & Wessel, 1990), and bi-harmonic splines (D. T. Sandwell, 1987). The latter belongs to a group of methods often called *radial basis functions* and is similar to the *thin-plate spline* (Franke, 1982). In these methods, the data are assumed to be represented by a linear combination of Green's functions,

$$d_i = \sum_{j=1}^{M} p_j G_j(\mathbf{x}_i, \mathbf{x}_j),$$

in which d_i is the ith datum, p_j is a scalar coefficient, G_j is a Green's function, and $\mathbf{x}_i$ and $\mathbf{x}_j$ are the position vectors for the datum and the point defining the Green's function, respectively. Interpolation is done by estimating the M p_j coefficients through linear least-squares and using them to predict data values at new locations on a grid. Essentially, these methods are linear models used for prediction. As such, many of the model selection and evaluation techniques used in machine learning can be applied to griding problems as well.

Verde is a Python library for gridding spatial data using different Green's functions. It differs from the radial basis functions in `scipy.interpolate` by providing an API inspired by scikit-learn (Pedregosa et al., 2011). The *Verde* API should be familiar to scikit-learn users but is tweaked to work with spatial data, which has Cartesian or geographic coordinates and multiple data components instead of an `X` feature matrix and `y` label vector. The library also includes more specialized Green's functions (D. T. Sandwell & Wessel, 2016), utilities for trend estimation and data decimation (which are often required prior to gridding (W. Smith & Wessel, 1990)), and more. Some of these interpolation and data processing methods already exist in the Generic Mapping Tools (GMT) (Wessel, Smith, Scharroo, Luis, & Wobbe, 2013), a command-line program popular in the Earth Sciences. However, there are no model selection tools in GMT and it can be

difficult to separate parts of the processing that are done internally by its modules. *Verde* is designed to be modular, easily extended, and integrated into the scientific Python ecosystem. It can be used to implement new interpolation methods by subclassing the `verde.base.BaseGridder` class, requiring only the implementation of the new Green's function. For example, it is currently being used to develop a method for interpolation of 3-component GPS data (Uieda, Sandwell, & Wessel, 2018).

Acknowledgements

I would like to thank Paul Wessel and David Hoese for helpful discussions and contributions to this project.

References

Franke, R. (1982). Smooth interpolation of scattered data by local thin plate splines. *Computers & Mathematics with Applications*, *8*(4), 273–281. doi:10.1016/0898-1221(82)90009-8

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., et al. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(Oct), 2825–2830.

Sandwell, D. T. (1987). Biharmonic spline interpolation of GEOS-3 and SEASAT altimeter data. *Geophysical Research Letters*, *14*(2), 139–142. doi:10.1029/GL014i002p00139

Sandwell, D. T., & Wessel, P. (2016). Interpolation of 2-D vector data using constraints from elasticity. *Geophysical Research Letters*, *43*(20), 2016GL070340. doi:10.1002/2016GL070340

Smith, W., & Wessel, P. (1990). Gridding with continuous curvature splines in tension. *Geophysics*, *55*(3), 293–305. doi:10.1190/1.1442837

Uieda, L., Sandwell, D., & Wessel, P. (2018). Presentation: Joint Interpolation of 3-component GPS Velocities Constrained by Elasticity. *figshare*. doi:10.6084/m9.figshare.6387467

Wessel, P., Smith, W. H. F., Scharroo, R., Luis, J., & Wobbe, F. (2013). Generic Mapping Tools: Improved Version Released. *Eos, Transactions American Geophysical Union*, *94*(45), 409–410. doi:10.1002/2013EO450001