The Journal of Open Source Software

MODULO: A Python toolbox for data-driven modal decomposition

R. Poletti [1,2], **L. Schena** [1,3], **D. Ninni** [4], **and M. A. Mendez** [1¶]

1 von Karman Insitute for Fluid Dynamics **2** University of Ghent, Belgium **3** Vrije Universiteit Brussel (VUB), Belgium **4** Politecnico di Bari, Italy ¶ Corresponding author

DOI: 10.21105/joss.06753

Software

- Review
- Repository
- Archive

Editor: Philip Cardiff

Reviewers:

- @mengaldo
- @AndreWeiner

Submitted: 09 January 2024
Published: 17 October 2024

License
Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License (CC BY 4.0).

Summary

Dimensionality reduction is an essential tool in processing large datasets, enabling data compression, pattern recognition, and reduced-order modeling. Many linear tools for dimensionality reduction have been developed in fluid mechanics, where they have been formulated to identify coherent structures and build reduced-order models of turbulent flows (Ranc et al., 2014). This work proposes a major upgrade of the software package MODULO (MODal mULtiscale pOd) (Ninni & Mendez, 2020), which was designed to perform Multiscale Proper Orthogonal Decomposition (mPOD) (Mendez et al., 2019). In addition to implementing the classic Fourier Transform (DFT) and Proper Orthogonal Decomposition (POD), MODULO now also allows for computing Dynamic Mode Decomposition (DMD) (Schmid, 2010) as well as the Spectral POD by Sieber et al. (2016), the Spectral POD by Towne et al. (2018) and a generalized kernel-based decomposition akin to kernel PCA (Mendez, 2023). All algorithms are wrapped in a 'SciKit'-like Python API, which allows computing all decompositions in one line of code. Documentation, exercises, and video tutorials are also provided to offer a primer on data drive modal analysis.

Statement of Need

As extensively illustrated in recent reviews (Mendez, 2023; Taira et al., 2020), all modal decompositions can be considered as special matrix factorizations. The matrix being factorized collects (many) snapshots (samples) of a high-dimensional variable. The factorization provides a basis for the matrix's column and row spaces to identify the most essential patterns (modes) according to a certain criterion. In what follows, we will refer to common terminologies in fluid dynamics. Nevertheless, it is worth stressing that these tools can be applied to any high-dimensional dataset to identify patterns and build reduced-order models (Mendez et al., 2019). In the common arrangement encountered in fluid dynamics, the basis for the column space is a set of 'spatial structures' while the basis for the row space is a set of 'temporal structures'. These are paired by a scalar, which defines their relative importance. The POD, closely related to Principal Component Analysis, yields modes with the highest energy (variance) content and, in addition, guarantees their orthonormality by construction. In the DFT, as implemented in MODULO, modes are defined to evolve as orthonormal complex exponential in time. This implies that the associated frequencies are integer multiples of a fundamental tone. The DMD generalizes the DFT by releasing the orthogonality constraint and considering complex frequencies, i.e., modes that can vanish or explode. Both the constraint of energy optimality and harmonic modes can lead to poor convergence and feature detection performances. This motivated the development of hybrid methods such as the Spectral POD by Towne et al. (2018), Spectral POD by Sieber et al. (2016), and Multiscale Proper Orthogonal Decomposition (mPOD) (Mendez et al., 2019). The first can be seen as an optimally averaged DMD, while the second combines POD and DFT with a filtering operation. Both SPODs

assume statistically stationary data and are designed to identify harmonic (or quasi-harmonic) modes. The mPOD combines POD with Multi-resolution Analysis (MRA) to provide optimal modes within a prescribed frequency band. The mPOD modes are thus spectrally less narrow than those obtained by the SPODs, but this allows for localizing them in time (i.e., potentially having compact support in time). Finally, recent developments in nonlinear methods such as kernel PCA and their applications to fluid dynamics (see Mendez (2023)) have motivated the interest in the connection between nonlinear methods and the most general Karhunen-Loeve expansion (KL). This generalizes the POD as the decomposition of data onto the eigenfunction of a kernel function (the POD being a KL for the case of linear kernel).

MODULO provides a unified tool to carry out different decompositions with a shared API. This simplifies comparing different techniques and streamlines their application to a given dataset (problem). In addition, it is the only package that includes the mPOD and the generalized KL with kernel functions interfacing with SciKit-learn. For decomposition-specific packages, we refer the reader to many excellent Python APIs that are available to compute the POD, DMD, and both SPODs, for example (Demo et al., 2018; Hatzissawidis & Sieber, 2023; Mengaldo & Maulik, 2021; Rogowski et al., 2024).

New Features

This manuscript accompanies MODULO version 2.0. This version features four new decompositions: the two SPODs, the DMD, and the general KL. It also allows for different approaches to computing the POD (interfacing with various SVD/EIG solvers from Scipy and SciKit-learn) and a first implementation for nonuniform grids. The memory-saving feature has been improved, and the software can now handle 3D decompositions.

Conclusions

MODULO is a versatile and user-friendly toolbox for data-driven modal decomposition. It provides a unified interface to various decomposition methods, allowing for a straightforward comparison and benchmarking. The package allows for modal decompositions in one line of code. It is also designed to handle large datasets via the so-called "memory saving" option and can handle nonuniformly sampled data.

Acknowledgements

R. Poletti and L. Schena are supported by Fonds Wetenschappelijk Onderzoek (FWO), grant numbers 1SD7823N and 1S75825N, respectively.

References

Demo, N., Tezzele, M., & Rozza, G. (2018). PyDMD: Python dynamic mode decomposition. *Journal of Open Source Software*, *3*(22), 530. https://doi.org/10.21105/joss.00530

Hatzissawidis, G., & Sieber, M. (2023). *Spectral proper orthogonal decomposition*. https://github.com/grigorishat/SPyOD.

Mendez, M. A. (2023). Linear and nonlinear dimensionality reduction from fluid mechanics to machine learning. *Measurement Science and Technology*, *34*, 042001. https://doi.org/10.1088/1361-6501/acaffe

Mendez, M. A., Balabane, M., & Buchlin, J.-M. (2019). Multi-scale proper orthogonal decomposition of complex fluid flows. *Journal of Fluid Mechanics*, *870*, 988–1036. https://doi.org/10.1017/jfm.2019.212

Mengaldo, G., & Maulik, R. (2021). PySPOD: A Python package for Spectral Proper Orthogonal Decomposition (SPOD). *Journal of Open Source Software*, *6*(60), 2862. https://doi.org/10.21105/joss.02862

Ninni, D., & Mendez, M. A. (2020). MODULO: A software for multiscale proper orthogonal decomposition of data. *SoftwareX*, *12*, 100622. https://doi.org/10.1016/j.softx.2020.100622

Ranc, N., Blanche, A., Ryckelynck, D., & Chrysochoos, A. (2014). POD preprocessing of IR thermal data to assess heat source distributions. *Experimental Mechanics*, *55*, 725–739. https://doi.org/10.1007/s11340-014-9858-2

Rogowski, M., Yeung, B. C. Y., Schmidt, O. T., Maulik, R., Dalcin, L., Parsani, M., & Mengaldo, G. (2024). Unlocking massively parallel spectral proper orthogonal decompositions in the PySPOD package. *Computer Physics Communications*, *302*, 109246. https://doi.org/10.1016/j.cpc.2024.109246

Schmid, P. J. (2010). Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics*, *656*, 5–28. https://doi.org/10.1017/S0022112010001217

Sieber, M., Paschereit, C. O., & Oberleithner, K. (2016). Spectral proper orthogonal decomposition. *Journal of Fluid Mechanics*, *792*, 798–828. https://doi.org/10.1017/jfm.2016.103

Taira, K., Hemati, M. S., Brunton, S. L., Sun, Y., Duraisamy, K., Bagheri, S., Dawson, S. T. M., & Yeh, C.-A. (2020). Modal analysis of fluid flows: Applications and outlook. *AIAA Journal*, *58*(3), 998–1022. https://doi.org/10.2514/1.J058462

Towne, A., Schmidt, O. T., & Colonius, T. (2018). Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis. *Journal of Fluid Mechanics*, *847*, 821–867. https://doi.org/10.1017/jfm.2018.283