The Journal of Open Source Software

MODULO: a Python toolbox for reduced order modeling

L. Schena[1,3], D. Ninni[2], and M. A. Mendez[1]

1 Department of Environmental and Applied Fluid Mechanics, the von Karman Institute (VKI) for Fluid Mechanics **2** Department of Aerospace Engineering, Politecnico di Bari **3** Department of Mechanical Engineering, Vrije Universiteit Brussels (VUB)

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

Summary

Dimensionality reduction is an essential tool in processing large datasets, enabling data compression, pattern recognition and reduced order modelling. 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 (Berkooz et al., 1993). 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) and the Spectral POD by (Towne et al., 2018). All algorithms are wrapped in a 'SciKit'-like Python API, which enables implementing all decompositions within a 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 (e.g., (Mendez, 2023), (Taira et al., 2020), all modal decompositions can be seen as special kinds of matrix factorizations. The matrix being factorized, herein denoted as D, collects (many) snapshots (samples) of a large dimensional variable. The factorization aims at providing a basis for the column and the row spaces of the matrix, with the goal of identifying the most essential patterns (modes) according to a certain criterion. 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, called amplitude or initial value, which defines their relative importance. In the POD, known as Principal Component Analysis in other fields, the modes are identified as the ones that have the largest amplitudes while remaining orthonormal both in space and time. In the classic 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 constraint of orthogonality and considering complex frequencies, i.e., modes that can potentially vanish or decay. Both the constraint of energy optimality and harmonic modes can lead to poor performances in terms of convergence and feature detection. This motivated the development of hybrid methods such as the the Spectral POD by (Towne et al., 2018), Spectral POD by (Sieber et al., 2016), Multiscale Proper Orthogonal Decomposition (mPOD)(Mendez et al., 2019). The first can be seen as an optimally averaged DMD while the second consists in bringing POD and DFT with the use of a filtering operation. Both the SPODs assume statistically stationary data and are mostly meant for identifying harmonic (or quasi-harmonic) modes. The mPOD on the other hand, consists in bridging POD with Multi-resolution Analysis (MRA), to provide modes that are optimal within a pre-scribed frequency band. The resulting mode are thus spectrally less narrow than those obtained in the SPODs, allowing also for their localization in time. While many excellent Python APIs are available to POD, DMD and both SPODs (see for example (Demo et al., 2018)) ((Brandt Belson, 2017), (Mengaldo & Maulik, 2021), (Hatzissaidis & Sieber, 2017)), MODULO is currently the only opensource tool for computing multiscale POD while still allowing to compute all the others.

New features

sDMD, SPOD

This release of `MODULO` includes the Spectral Proper Orthogonal Decomposition (SPOD) of Towne et al. (Towne et al., 2018) and the one proposed by Sieber et al. (Sieber et al., 2016). Even though these go with the same name, their algorithms are noticeably different. The Towne's methodology splits the dataset $D(\mathbf{x}, t)$ in n_b equal portions, scaled in time. Then, the DFT must computed on each block, originating a tensor of size $(n_s \times n_f \times n_b)$, where n_s are the degree of freedom of the dataset - i.e. spatial points, n_f depends on the frequency bins of the DFT and finally n_b are the different blocks assembled before, depending on the time horizon of the measurement. Finally, the classical POD is performed on each frequency bin, originating $n_f \times rank(D)$ modes. On the other hand, the Spectral approach of Sieber et al. consists in filtering the correlation matrix $\mathbf{K} \in \mathbb{R}^{n_t \times n_t}$, computed as in a normal POD decomposition. In particular, a pass-bass filter is applied on its diagonals which enforces the POD to be closer to a DFT. In fact, for a stationary process the $\mathbf{K}$ matrix would become a Toeplitz circulant matrix, which eigenvalues are indeed the DFT modes. After being filtered, then the diagonalization proceeds as for the POD.

Conclusion

700 characters: what is the goal of modulo?

Acknowledgments

See if needed to acknowledge something/someone ((**miguel?**))

References

Berkooz, G., Holmes, P., & Lumley, J. L. (1993). The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows. *Annual Review of Fluid Mechanics*, *25*(1), 539–575. https://doi.org/10.1146/annurev.fl.25.010193.002543

Brandt Belson, C. R., Jonathan Tu. (2017). *Modred*. https://pypi.org/project/modred/.

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

Hatzissaidis, G., & Sieber, M. (2017). *Modred*. https://pypi.org/project/modred/.

Mendez, M. A. (2023). Generalized and multiscale modal analysis. In M. A. Mendez, A. Ianiro, B. R. Noack, & S. L. E. Brunton (Eds.), *Data-driven fluid mechanics: Combining first principles and machine learning* (pp. 153–181). Cambridge University Press. https://doi.org/10.1017/9781108896214.013

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/https://doi.org/10.1016/j.softx.2020.100622

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