Physics-preserving Projections for Compression of Data and Dynamical Systems

Patrick Buchfink[a] and Bernard Haasdonk[a]

[a] *Institute of Applied Analysis and Numerical Simulation (IANS), University of Stuttgart, Germany*

Mathematical models for physical phenomena possess in many cases a certain structure. Hamiltonian systems are an example for such structured systems which preserve physical quantities, e.g. the Hamiltonian (function) over time. Thus, these systems are well-suited to model non-dissipative phenomena.

When faced with data from such systems, e.g. measurements or field solutions of numerical computations, it is important to take the structure into account as prior knowledge in the data analysis task. In computational mathematics, preservation of such structures shows great improvements in stability and accuracy e.g. for numerical integration [2] or model reduction [3]. Similar results can be expected for data analysis tasks such as feature extraction, filtering, classification, regression or data reconstruction.

The underlying structure of a Hamiltonian system is the so-called symplectic geometry. For finite-dimensional systems, the Hamiltonian system is described by the triple $(\mathbb{V}, \omega_{2n}, \mathcal{H})$ with

1. a (necessarily even-dimensional) phase-space $\mathbb{V}$ which we identify with $\mathbb{V} \cong \mathbb{R}^{2n}$,
2. a symplectic form $\omega_{2n} : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ which is a skew-symmetric, non-degenerate bilinear form,
3. a smooth Hamiltonian function $\mathcal{H} : \mathbb{V} \to \mathbb{R}$.

The tuple $(\mathbb{V}, \omega_{2n})$ is called symplectic (linear) vector space. In the canonical case, the symplectic form is given by

$$\omega_{2n}(\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{v}^{\mathsf{T}} \mathbb{J}_{2n} \boldsymbol{w}, \qquad \mathbb{J}_{2n} = \begin{bmatrix} \mathbf{0}_n & \boldsymbol{I}_n \\ -\boldsymbol{I}_n & \mathbf{0}_n \end{bmatrix}$$

where $\mathbb{J}_{2n}$ is the so-called (canonical) Poisson matrix composed of identity matrices $\boldsymbol{I}_n \in \mathbb{R}^{n \times n}$ and matrices with all-zeros $\mathbf{0}_n \in \mathbb{R}^{n \times n}$.

We present two data-based methods to extract

(a) the most important $2k$-dimensional orthogonal, symplectic subspace $(\mathbb{V}_{\boldsymbol{B}}, \langle \cdot, \cdot \rangle, \omega_{2k})$ with the canonical scalar product $\langle \cdot, \cdot \rangle : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ or

(b) a $2k$-dimensional (in general) non-orthogonal but symplectic subspace $(\mathbb{V}_{\boldsymbol{B}}, \omega_{2k})$

spanned by a given data set. Both methods are findings in the field of model reduction [3, 1] but should be considered in a broader view of physics-preserving projections for data analysis.

To this end, we stack our $m \in \mathbb{N}$ data samples of even dimension $2n \in \mathbb{N}$ as vectors $\boldsymbol{b}_i \in \mathbb{R}^{2n}$, $i = 1, \ldots, m$, in the columns of $\boldsymbol{B} = [\boldsymbol{b}_1, \ldots, \boldsymbol{b}_m] \in \mathbb{R}^{2n \times m}$. To derive the optimal orthogonal, symplectic subspace (a), a complexification of the data matrix $\boldsymbol{B}$ and the Complex SVD are used following [3]. The result is a minimizer of the optimization problem

$$\underset{\boldsymbol{V} \in \mathbb{R}^{2n \times 2k}}{\text{minimize}} \sum_{i=1}^{m} \left\| (\boldsymbol{I}_{2n} - \boldsymbol{V} \underbrace{\mathbb{J}_{2k}^{\mathsf{T}} \boldsymbol{V}^{\mathsf{T}} \mathbb{J}_{2n}}_{=: \boldsymbol{V}^{+}}) \boldsymbol{b}_i \right\|_2^2 \qquad \text{such that} \quad \boldsymbol{V}^{\mathsf{T}} \boldsymbol{V} = \boldsymbol{I}_{2k} \quad \text{and} \quad \boldsymbol{V}^{+} \boldsymbol{V} = \boldsymbol{I}_{2k}.$$

which is the sum of the squared errors of the residuals $(\boldsymbol{I}_{2n} - \boldsymbol{V}\boldsymbol{V}^{+})\boldsymbol{b}_i$ of the data sample $\boldsymbol{b}_i$ with respect to the symplectic projection $P = \boldsymbol{V}\boldsymbol{V}^{+}$ onto an orthogonal, symplectic subspace. We emphasize that $\boldsymbol{V}^{+} = \mathbb{J}_{2k}^{\mathsf{T}} \boldsymbol{V}^{\mathsf{T}} \mathbb{J}_{2n}$ is the so-called symplectic inverse which is in general not equal to the Moore-Penrose pseudo inverse which occasionally uses the same notation $(\cdot)^{+}$.

To derive a non-orthogonal, symplectic subspace (b), we use an SVD-like decomposition [4, 5] of the data matrix $\boldsymbol{B}$ which reads

$$\boldsymbol{B} = \boldsymbol{S}\boldsymbol{D}\boldsymbol{Q}^{\mathsf{T}}, \qquad \begin{cases} \boldsymbol{S} \in \mathbb{R}^{2n\times 2n} & \text{symplectic matrix} \\ \boldsymbol{D} \in \mathbb{R}^{2n\times m} & \\ \boldsymbol{Q} \in \mathbb{R}^{m\times m} & \text{orthogonal matrix} \end{cases}, \qquad \boldsymbol{D} = \begin{array}{c} \begin{array}{cccc} p & q & p & m-2p-q \end{array} \\ \left(\begin{array}{cccc} \boldsymbol{\Sigma}_{\mathrm{s}} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_q & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{\Sigma}_{\mathrm{s}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{array}\right) \end{array} \begin{array}{c} \\ p \\ q \\ n-p-q \\ p \\ n-p \end{array},$$

with the so-called symplectic singular values $\boldsymbol{\Sigma}_{\mathrm{s}} = \operatorname{diag}(\sigma_1^{\mathrm{s}}, \dots, \sigma_p^{\mathrm{s}}) \in \mathbb{R}^{p\times p}$, $\sigma_i^{\mathrm{s}} > 0$ for $1 \le i \le p$ where $p, q \in \mathbb{N}$ contain information of the "degree of symplecticity" of the image of $\boldsymbol{B}$. The approach [1] then follows the central idea of the Principal Component Analysis (PCA) (also known as Proper Orthogonal Decomposition (POD)) which is used to derive the most important $2k$-dimensional orthogonal subspace of the image of $\boldsymbol{B}$: the extraction of the symplectic subspace relies on the truncation of the symplectic singular vectors $\boldsymbol{S} = [\boldsymbol{s}_1, \dots, \boldsymbol{s}_{2n}]$ of an SVD-like decomposition instead of the left-singular vectors $\boldsymbol{U} = [\boldsymbol{u}_1, \dots, \boldsymbol{u}_{2n}]$ of an SVD $\boldsymbol{B} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^{\mathsf{T}}$ (which is done in PCA).

As application of both methods (a) and (b), we consider three scenarios:

1. we compress simulation data obtained from numerical simulations and give error bounds for the compression,
2. we compress/reduce the dimension of a Hamiltonian system using symplectic model reduction,
3. we estimate the state for a given set of measurements which is a data reconstruction task.

References

[1] P. Buchfink, A. Bhatt, and B. Haasdonk. Symplectic Model Order Reduction with Non-Orthonormal Bases. *Math. Comput. Appl.*, 24(2), 2019. doi: 10.3390/mca24020043.

[2] E. Hairer, G. Wanner, and C. Lubich. *Geometric Numerical Integration*. Springer, Berlin, Heidelberg, 2006. ISBN 978-3-540-30666-5. doi: 10.1007/3-540-30666-8.

[3] L. Peng and K. Mohseni. Symplectic Model Reduction of Hamiltonian Systems. *SIAM J. Sci. Comput.*, 38(1):A1–A27, 2016. doi: 10.1137/140978922.

[4] H. Xu. An SVD-like matrix decomposition and its applications. *Linear Algebra Appl.*, 368:1 – 24, 2003. doi: 10.1016/S0024-3795(03)00370-7.

[5] H. Xu. A Numerical Method for Computing an SVD-like Decomposition. *SIAM J. Matrix Anal. Appl.*, 26(4):1058–1082, 2005. doi: 10.1137/S0895479802410529.