Interpolation on the manifold of fixed-rank positive-semidefinite matrices for parametric model order reduction: preliminary results

Estelle Massart[1], Pierre-Yves Gousenbourger[1], Nguyen Thanh Son[1,2], Tatjana Stykel[3], P.-A. Absil[1] *

1- ICTEAM Institute, UCLouvain, Louvain-la-Neuve, Belgium

2- Dept. Mathematics and Informatics, TNUS, Thai Nguyen, Vietnam

3- Institute of Mathematics, University of Augsburg, Augsburg, Germany

Abstract. We present several interpolation schemes on the manifold of fixed-rank positive-semidefinite (PSD) matrices. We explain how these techniques can be used for model order reduction of parameterized linear dynamical systems, and obtain preliminary results on an application.

1 Introduction

Model order reduction (MOR) is a well known tool to simulate large-scale systems [1]. Its purpose is to compute a model of smaller order (by design, faster to simulate) that represents accurately enough the behavior of the full-order system. The reduced-order model is usually obtained via projection-based methods such as balanced truncation (further described in Section 2), proper orthogonal decomposition, Krylov subspace-based moments matching or $\mathcal{H}_2$-norm optimization [1]. Those methods provide two matrices V_{Proj} and W_{Proj} that project the large-scale system into a smaller state space.

In many cases, the system depends on parameters representing, e.g., physical, material or environmental properties [2]. *Parametric* model order reduction (PMOR) computes a parameterized reduced model that approximates the behavior of the full-order system in a given parameter range.

A classical technique of PMOR is made of two steps. The reduced-order models are obtained for a subset of parameter values using MOR (offline step). They are used to recover the reduced models for other parameter values (online step), by interpolation. The interpolation is done on representatives of the models, e.g., the projection matrices V_{Proj} and W_{Proj}, the reduced-system matrices, or the reduced transfer functions. We refer the reader to [3] for more information about PMOR.

In this paper, we consider balanced truncation as the offline step of the PMOR. The most computationally expensive part of this algorithm is the computation of a low-rank approximation of the positive-semidefinite (PSD) solutions P and Q to a pair of Lyapunov equations. We propose to compute those

*Acknowledgments: this work was supported by (i) the Fonds de la Recherche Scientifique – FNRS and the Fonds Wetenschappelijk Onderzoek – Vlaanderen under EOS Project no 30468160 and (ii) "Communauté française de Belgique - Actions de Recherche Concertées" (contract ARC 14/19-060).

low-rank approximations for a few values of the parameters (offline). Then, for other values of the parameters, the precomputed solutions are interpolated (online) on the manifold $\mathcal{S}_+(p,n)$ of $n \times n$ positive-semidefinite matrices of rank p.

We consider different interpolation schemes on $\mathcal{S}_+(p,n)$. Interpolation on manifolds has been a hot topic for a few years (see [4, 5] and references therein) and has been recently applied to $\mathcal{S}_+(p,n)$ in [6, 7]. We compare several interpolation algorithms regarding their ability to recover the solutions to the Lyapunov equations associated with (intermediary) parameter values. We observe numerically that the interpolation method by blended cubic splines recently proposed in [5] seems to be the best choice, among the five methods considered. They also indicate that the computation time required by the interpolation algorithms is several orders of magnitude lower than the time needed to run the ADI solver developed in [8].

The paper is organized as follows. In Section 2, we explain in more detail how interpolation on $\mathcal{S}_+(p,n)$ can be used for PMOR. The geometric structure of $\mathcal{S}_+(p,n)$ and the different interpolation methods are presented in Section 3. Numerical results are shown in Section 4.

2 Parametric model order reduction

We briefly recall the theory on PMOR and on the balanced truncation method (see [1] for more information). We consider an asymptotically stable linear parameterized system:

$$\begin{aligned} E(\mu)\dot{x}(t,\mu) &= A(\mu)x(t,\mu) + B(\mu)u(t), \\ y(t,\mu) &= C(\mu)x(t,\mu), \end{aligned} \tag{1}$$

with $\mu \in [\alpha, \beta]$, a parameter representing, e.g., physical, material or environmental properties, $E(\mu) \in \mathbb{R}^{n\times n}$ nonsingular, $A(\mu) \in \mathbb{R}^{n\times n}$, $B(\mu) \in \mathbb{R}^{n\times m}$, $C(\mu) \in \mathbb{R}^{s\times n}$, and $s, m \ll n$. The vectors $x(t,\mu), u(t)$ and $y(t,\mu)$ are respectively the state, input and output vectors of the system. The goal of PMOR is to approximate system (1) with a parametric reduced-order model

$$\begin{aligned} \tilde{E}(\mu)\dot{\tilde{x}}(t,\mu) &= \tilde{A}(\mu)\tilde{x}(t,\mu) + \tilde{B}(\mu)u(t), \\ \tilde{y}(t,\mu) &= \tilde{C}(\mu)\tilde{x}(t,\mu), \end{aligned} \tag{2}$$

where $\tilde{E}(\mu)$, $\tilde{A}(\mu) \in \mathbb{R}^{r\times r}$, $\tilde{B}(\mu) \in \mathbb{R}^{r\times m}$, $\tilde{C}(\mu) \in \mathbb{R}^{s\times r}$ and $r \ll n$ (we will now omit the μ-dependency for readability).

Balanced truncation is a three-step method to compute the reduced-order model. First, one has to find the low-rank approximate solutions $P = XX^\top$ and $Q = YY^\top$ of the Lyapunov equations

$$EPA^\top + APE^\top = -BB^\top, \qquad E^\top QA + A^\top QE = -C^\top C, \tag{3}$$

where $X \in \mathbb{R}^{n\times k_X}$, $Y \in \mathbb{R}^{n\times k_Y}$, in which k_X and k_Y depend on μ as well. In this paper, we resort to the low-rank ADI solver [8]. Secondly, one computes the singular value decomposition $Y^\top EX = U\Sigma V^\top$. This SVD is truncated to a

given rank r such that $\tilde{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$ is composed of the r first largest singular values $\{\sigma_i\}_{i=1}^r$, and $\tilde{U}$ and $\tilde{V}$ are respectively the truncation of U and V to their r first columns. The projection matrices of the reduced-order model are then given by

$$W_{\mathrm{Proj}} = Y\tilde{U}\tilde{\Sigma}^{-1/2}, \qquad V_{\mathrm{Proj}} = X\tilde{V}\tilde{\Sigma}^{-1/2}. \tag{4}$$

Finally, the reduced model (2) is obtained by projection: $\tilde{E} = W_{\mathrm{Proj}}^\top E V_{\mathrm{Proj}}$, $\tilde{A} = W_{\mathrm{Proj}}^\top A V_{\mathrm{Proj}}$, $\tilde{B} = W_{\mathrm{Proj}}^\top B$, and $\tilde{C} = C V_{\mathrm{Proj}}$.

Note that the truncation rank r is related to the error between the transfer function $\tilde{H}$ of the reduced model (2) and H, the one of the full-order model (1), such that it can be chosen relatively to a given tolerance ϵ as

$$\|H - \tilde{H}\|_{\mathcal{H}_\infty} \leq 2\big(\sigma_{r+1} + \cdots + \sigma_{\min(k_X, k_Y)}\big) < \epsilon.$$

For a transfer function G, $\|G\|_{\mathcal{H}_\infty}$ is the supremum, over the frequencies, of the magnitude of G. Instead of computing the solutions P and Q of (3) for each value of μ, we propose to precompute solutions P_i and Q_i, associated to some parameter values μ_i, $i = 1, \ldots, N$, and to recover P (respectively Q) by interpolation of the P_i (respectively the Q_i). The precomputed solutions P_i and Q_i are PSD, with rank k_{X_i} and k_{Y_i}, respectively. On data with such properties, interpolation methods on manifolds were shown to be valuable [6, 7]. However, the set of all PSD matrices is not a manifold. Therefore, we truncate the rank of the matrices to a given rank p, such that they belong to $\mathcal{S}_+(p, n)$. For instance, to interpolate the P_i, we choose $p = \min(\{k_{X_i}\}_{i=1}^N)$. This step induces a loss of information, which turns out to be mild in the numerical experiments presented in this paper. We define $\tilde{P}_i := \tilde{X}_i \tilde{X}_i^\top$, where $\tilde{X}_i$ is made of the p first columns of X_i. By the design of the low-rank ADI method, these columns contain the dominant information of the low-rank solution. The same reasoning can be used for the interpolation of the Q_i, choosing then $p := \min(\{k_{Y_i}\}_{i=1}^N)$. The next section presents some interpolation algorithms on $\mathcal{S}_+(p, n)$.

Once the interpolant $\tilde{P}$ and $\tilde{Q}$, associated to μ, are known, the construction of the reduced system (2) is considerably cheaper than the resolution of the Lyapunov equations.

3 Interpolation on $\mathcal{S}_+(p, n)$

Roughly speaking, a Riemannian manifold $\mathcal{M}$ is a nonlinear space that can be approximated around any point $x \in \mathcal{M}$ by a Euclidean tangent space $T_x\mathcal{M}$. Two main tools are usually required to perform efficient computations on $\mathcal{M}$: the logarithm map $\log_x y$ to map a given point $y \in \mathcal{M}$ to the tangent space $T_x\mathcal{M}$, and the exponential map $\exp_x \eta$ (the reverse operation) to map a tangent vector $\eta \in T_x\mathcal{M}$ towards $\mathcal{M}$. The curve $\gamma(\cdot; x, y) : [0, 1] \to \mathcal{M}$ of zero acceleration between x and y is called geodesic and can be computed as $\gamma(t; x, y) = \exp_x(t \log_x y)$.

3.1 The manifold $\mathcal{S}_+(p,n)$

The set $\mathcal{S}_+(p,n)$ is a non-Euclidean space that can be endowed in several ways with a Riemannian manifold structure [9, §7]. We refer here to the approach developed in [10], in which $\mathcal{S}_+(p,n)$ is seen as a quotient manifold $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$, with $\mathbb{R}_*^{n\times p}$ the set of full-rank $n\times p$ matrices, and $\mathcal{O}_p$ the orthogonal group in dimension p. In this representation, any PSD matrix S is factorized as $S=YY^\top$, with $Y\in\mathbb{R}_*^{n\times p}$. The quotient structure comes from the fact that $\hat{Y}\hat{Y}^\top = YY^\top$ if and only if $\hat{Y}=YQ$ with $Q\in\mathcal{O}_p$. This representation allows to perform computations directly on the Y-factors, instead of on the full PSD matrices. This is particularly suitable to our case, as to define V_{Proj} and W_{Proj}, we only require the factors $\tilde{X},\tilde{Y}$ of the PSD matrices $\tilde{P}:=\tilde{X}\tilde{X}^\top$ and $\tilde{Q}:=\tilde{Y}\tilde{Y}^\top$.

We endow the space $\mathbb{R}_*^{n\times p}$ with the Euclidean metric. In the Y-representation, the computation of the exponential map translates simply into a sum, while the logarithm map requires only to compute a sum and the polar decomposition of a $p\times p$ matrix. However, caution should be taken, as exp and log are not everywhere defined, and moreover $\log\circ\exp$ is the identity on a further restricted domain (denoted by $\mathcal{M}_Y$ in [10]). We make the standing assumptions that exp and log need only be evaluated at points where these issues do not arise. We finally define the projection operator $\Pi:\mathcal{S}_+(k,n)\to\mathcal{S}_+(p,n)$, that returns, for a PSD matrix of rank $k\geq p$, the (generally unique) closest PSD matrix of rank p.

3.2 Interpolation methods on manifolds

We now briefly present several interpolation methods on $\mathcal{S}_+(p,n)$. The curves proposed interpolate a set of N points $P_1,\dots,P_N$ associated with parameters $\mu_1<\cdots<\mu_N$.

Interpolation in the ambient space. A first approach consists in running Euclidean interpolation algorithms in the ambient space $\mathbb{R}^{n\times n}$ (i.e., disregarding the fact that the data lie on a manifold). The result is then projected back on $\mathcal{S}_+(p,n)$, using the projection operator Π. We consider here the curve $\mathbf{P}_{\text{AS}}^{\text{lin}}(\mu)$, corresponding to piecewise **lin**ear interpolation in the **A**mbient **S**pace $\mathbb{R}^{n\times n}$.

Interpolation on one tangent space. A second approach consists in a three-step procedure: (i) all the data points are mapped to the tangent space based at an arbitrary point P_{ref}, using the $\log_{P_{\text{ref}}}$ map; (ii) Euclidean interpolation is performed on those points; (iii) the result is finally mapped back on the manifold using the $\exp_{P_{\text{ref}}}$ map. The curve $\mathbf{P}_{\text{TS}}^{\text{cub}}(\mu)$ is obtained with this strategy, interpolation in the **T**angent **S**pace being done with **cub**ic splines. We consider two possible choices for the tangent space: $P_{\text{ref}}=P_1$ or $P_{\text{ref}}=P_{\lfloor (N+1)/2\rfloor}$.

Blended cubic splines and piecewise geodesic splines. More advanced algorithms on manifolds do not perform all the computations on one single tangent space (which can generate significant deviations when some points are far from the root of the tangent space). The piecewise geodesic spline $\mathbf{P}_{\text{PG}}(\mu)$ consists in a

concatenation of geodesics between two consecutive data points. The blended cubic spline $\mathbf{P}_{\text{Blend}}(\mu)$ is constructed as a C^1 composite curve

$$\mathbf{P}_{\text{Blend}} : [0,1] \to \mathcal{M} : f_i(\mu - \mu_i),\ i = 0, \ldots, N-1,$$

made of so-called blended functions f_i. These blended functions are obtained by blending together interpolating curves from different tangent spaces. In this case, f_i is the weighted mean of two curves $\mathbf{P}^{\text{cub}}_{\text{TS}}(\mu)$ computed respectively on the tangent spaces based at d_i and d_{i+1}. The weights are chosen such that $\mathbf{P}_{\text{Blend}}(\mu)$ is a C^1 interpolating curve. Blended cubic splines were proposed in [5] as alternatives to Bézier curves on manifolds when conditions on the data points were not satisfied to ensure a correct interpolation.

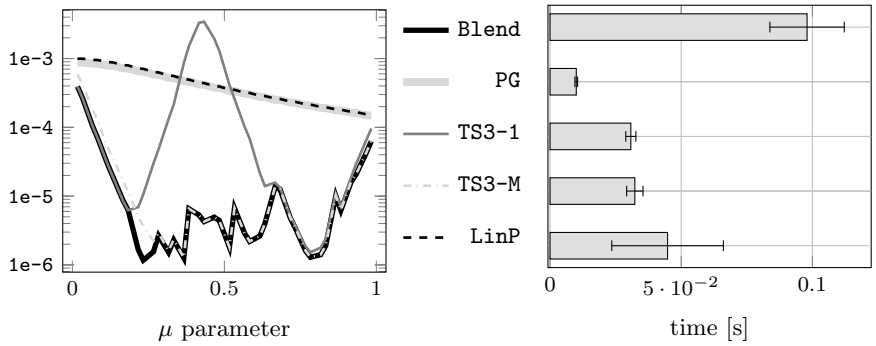
4 Numerical results

In our experiments, the full-order model (1) is the 1-parameter anomometer [2] with $n = 29008$, and $\mu \in [0,1]$ is the fluid velocity. We run the low-rank ADI method [8] to solve the Lyapunov equation (3) defining P (with tolerance $\epsilon = 10^{-8}$), and obtain a training set $\{P_1^{\text{ADI}}, \ldots, P_{21}^{\text{ADI}}\}$, with $\{\mu_1, \ldots, \mu_{21}\} = \{0, 0.05, \ldots, 1\}$. The rank of the solutions returned by the solver varies from 25 to 39, so that $p = 25$. We compare the ability of the interpolation methods to recover P^{ADI} at test values of μ.

Figure 1 (left) shows the relative error $\text{E}_{\text{rel}}(\mu)$ between the predicted matrices $P(\mu)$ and the ADI solutions $P^{\text{ADI}}(\mu)$, truncated to a rank $p = 25$:

$$\text{E}_{\text{rel}}(\mu) = \frac{||P(\mu) - P^{\text{ADI}}(\mu)||_{\text{F}}}{||P^{\text{ADI}}(\mu))||_{\text{F}}}.$$

The test set is made of 40 points, and has no intersection with the training set, for which E_{rel} is zero. The best trade-off between computation time and accuracy seems to be the cubic interpolation in the tangent space $\mathbf{P}^{\text{cub}}_{\text{TS}}(\mu)$, choosing P_{ref} as the midpoint of the data set (`TS3-M`). However, this method is sensitive to the choice of the tangent space, as setting $P_{\text{ref}} = P_1$ leads to significantly larger errors (`TS3-1`). The blended curve, which intrinsically combines several tangent spaces, does not present this drawback, and reaches the same accuracy as `TS3-M`. The fact that $\mathbf{P}^{\text{cub}}_{\text{TS}}(\mu)$ and $\mathbf{P}_{\text{Blend}}(\mu)$ are similar when the tangent space is based at the midpoint of the data set indicates that the curvature of the manifold is small around the data points considered. For the same reason, the piecewise linear interpolation in the ambient space $\mathbf{P}^{\text{lin}}_{\text{AS}}(\mu)$ (`LinP`) is almost as accurate as the piecewise geodesic spline $\mathbf{P}_{\text{PG}}(\mu)$ (`PG`). Figure 1 (right) compares the computation times required to obtain the value of the different curves at one arbitrary value of μ (all the computations that were independent of μ were done offline and not considered in those timings). The curve $\mathbf{P}_{\text{Blend}}(\mu)$ is the most expensive one. However, those values have to be compared with the average time to compute the matrix $P^{\text{ADI}}(\mu)$ using the ADI solver (i.e., without interpolation), which is here around 10s. Extension to more parameters and comparison of our PMOR procedure (based on interpolation on $\mathcal{S}_+(p,n)$) with state-of-the-art PMOR methods, are left for future work.



(a) Relative error E_{rel} **(b)** Average interpolation time

Figure 1: Comparison of the five interpolation methods considered.

References

[1] Athanasios C. Antoulas. *Approximation of Large-Scale Dynamical Systems*. SIAM, Philadelphia, PA, 2005.

[2] MOR Wiki – Model Order Reduction Wiki. http://morwiki.mpi-magdeburg.mpg.de/morwiki/.

[3] Peter Benner, Serkan Gugercin, and Karen Willcox. A survey of projection-based model reduction methods for parametric dynamical systems. *SIAM Review*, 57(4):483–531, 2015.

[4] Luís Machado, Fatima Silva Leite, and Maria Teresa T. Monteiro. Path planning trajectories in fluid environments. In *2018 13th APCA International Conference on Control and Soft Computing (CONTROLO)*, volume 1, pages 219–223, 2018.

[5] Pierre-Yves Gousenbourger, Estelle Massart, and P.-A. Absil. Data fitting on manifolds with composite Bézier-like curves and blended cubic splines. *Journal of Mathematical Imaging and Vision*, 2018. Preprint: https://sites.uclouvain.be/absil/2018.04.

[6] Pierre-Yves Gousenbourger, Estelle Massart, Antoni Musolas, P.-A. Absil, Laurent Jacques, Julien M. Hendrickx, and Youssef Marzouk. Piecewise-Bézier C^1 smoothing on manifolds with application to wind field estimation. In *ESANN2017*, pages 305–310. Springer, 2017.

[7] Anis Kacem, Mohamed Daoudi, Boulbaba Ben Amor, Stefano Berretti, and Juan Carlos Alvarez-Paiva. A Novel Geometric Framework on Gram Matrix Trajectories for Human Behavior Understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)*, 2018.

[8] Jing-Rebecca Li and Jacob White. Low rank solution of Lyapunov equations. *SIAM J. Matrix Anal. Appl.*, 24(1):260–280, 2002.

[9] Bart Vandereycken, P.-A. Absil, and Stefan Vandewalle. A Riemannian geometry with complete geodesics for the set of positive semidefinite matrices of fixed rank. *IMA Journal of Numerical Analysis*, 33(2):481–514, 2013.

[10] Estelle Massart and P.-A. Absil. Quotient geometry with simple geodesics for the manifold of fixed-rank positive-semidefinite matrices. Technical Report UCL-INMA-2018.06, UCLouvain, November 2018. Preprint: http://sites.uclouvain.be/absil/2018.06.