Chapter 2

Nonlinear Model Reduction with POD-DEIM

This Chapter presents a model reduction method for nonlinear ordinary differential equation (ODEs). First a general formulation of the problem and its reduction via Proper Orthogonal Decomposition (POD) with Galerkin projection continued with a discussion of its complexity issue in the nonlinear part. To resolve this issue the Discrete Empirical Interpolation Method (DEIM) is introduced in 2.3. This method replace the orthogonal projection of POD in the nonlinear part with an interpolating projection of DEIM. The DEIM was fist introduced by Saifon Chaturantabut in [1], where also the algorithm and the error bound analysis presented in 2.3 comes from.

2.1 Problem Formulation

Consider a system of nonlinear differential equations of the form

$$\frac{d}{dt}y(t) = Ay(t) + q(y(t)), \tag{2.1}$$

where $t \in [0.T]$ denotes the time, $y(t) = [y_1(t), ..., y_n(t)]^T \in \mathbb{R}^n$ is a vector of a state with initial condition $y(0) = y_0 \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$ is the discrete approximation of the linear differential operator and $q : \mathbb{R}^n \longmapsto \mathbb{R}^n$ a nonlinear vector-based function. The complexity, both time and space, to solve the system is mostly depended of the dimension of the system n, which could be very large, if high accuracy is required, so to reduce the complexity, we want to reduce the dimension.

In order to do this the system will be projected on a subspace spanned in $\mathbb{R}^n$ by a reduced basis of a dimension $k \ll n$. These projection-based techniques are commonly used for constructing a reduced-order system, that approximates the original system in the subspace. Here a Galerkin projection is used as the means for dimension reduction. For that let $V \in \mathbb{R}^{n \times k}$ a matrix whose columns are the orthonormal vectors of the reduced basis. Then projecting the system 2.1 onto V by replacing $y(t)$ with $V\hat{y}(t)$ where $\hat{y}(t) \in \mathbb{R}^k$, the reduced system is of the form

$$\frac{d}{dt}\hat{y}(t) = V^T A V \hat{y}(t) + V^T q(V\hat{y}(t)). \tag{2.2}$$

The quality of the approximation is clearly affected by the choice of the reduced basis. POD constructs a set of basis vectors from a singular value decomposition (SVD) of snapshots

$$S = y_1, ..., y_m, \tag{2.3}$$

which are samples of trajectories $y(\cdot)$ for a particular set of parameters, boundary condition and other system inputs. After a reduced model has been constructed from this basis, it may be used to compute approximate solutions for different initial conditions and parameter settings. If the snapshots are diverse enough it is expected that the approximate solution is near to the high dimensional one. The POD basis is optimal in the sense that a approximation error in relation to the snapshots is minimized. Therefore the POD approach is used here for constructing the basis.

In the equation 2.2 the linear part is already reduced in dimension, because the precomputation of $\hat{A} = V^T A V \in \mathbb{R}^{k \times k}$ supplies a system of dimension $k \ll n$. But the evaluation of the nonlinear function $q(y(t))$ is still in the dimension of n

$$\frac{d}{dt}\hat{y}(t) = \underbrace{\hat{A}}_{k \times k}\hat{y}(t) + V^T q(\underbrace{V\hat{y}(t)}_{n}). \tag{2.4}$$

To reduce also the nonlinear part, the DEIM will be used and this provides a approximation $\hat{q}(y(t))$ of the nonlinear function by projecting the function $q(y(t))$ onto a $m \ll n$ dimensional subspace spanned by another basis $U \in \mathbb{R}^{n\times m}$, constructed out of snapshots of the nonlinear function with the POD-method. With coefficients $c(y(t))$ the approximation is of the form

$$q(y(t)) \approx \hat{q}(y(t)) = Uc(y(t)) \tag{2.5}$$

and the form of the reduced system is

$$\frac{d}{dt}\hat{y}(t) = \underbrace{\hat{A}}_{k\times k}\hat{y}(t) + \underbrace{V^T U}_{k\times m} c(V\hat{y}(t)). \tag{2.6}$$

The system of 2.1 is now completely reduced in dimension, the linear part is reduced to $k \ll n$ and the nonlinear part to $m \ll n$.

How the coefficients $c(y(t))$ will be determined, will be explained in Chapter 2.3, along with an explanation of the DEIM-algorithm and an short error analysis. How the two basis are constructing using the POD-method will be explained in then next Chapter.

2.2 Proper Orthogonal Decomposition (POD)

Consider a set of snapshots $S = \{y_1, ..., y_{n_s}\} \subset \mathbb{R}^n$ and the corresponding snapshots matrix $Y = [y_1, ..., y_{n_s}] \in \mathbb{R}^{n\times n_s}$. The POD method constructs an orthonormal basis in the space spanned by S. Let $r = rank\{Y\}$, then there exist $\{v_i\}_i^k \subset \mathbb{R}^n$ orthonormal basis vectors, for a $k \leq r$.

Definition 1 ***(SVD):*** *Let $\mathbb{A} \in \mathbb{R}^{m\times n}$ a matrix of rank r, then there exist two orthogonal matrices*

$$U = [\varphi_1|...|\varphi_n] \in \mathbb{R}^{m\times m}, \qquad Z = [\Psi_1|...|\Psi_m] \in \mathbb{R}^{n\times n}$$

such that

$$\mathbb{A} = U\Sigma Z^T,\ mit\ \Sigma = diag(\sigma_1, ..., \sigma_p) \in \mathbb{R}^{m\times n} \tag{2.7}$$

and $\sigma_1 \geq ... \geq \sigma_r > 0$, $\mathbb{A}\Psi_i = \sigma_i\varphi_i$ *and* $\mathbb{A}^T\varphi_i = \sigma_i\Psi_i$ *with* $i = 1, ..., r$.

The SVD of $Y = U\Sigma Z^T$ with orthogonal matrices $U = [\varphi_1|...|\varphi_n] \in \mathbb{R}^{n\times n}$ and $Z = [\Psi_1|...|\Psi_m] \in \mathbb{R}^{m\times m}$ and the diagonal matrix $\Sigma = diag(\sigma_1 \geq .... \geq \sigma_r)$ with singular values $\sigma_1 \geq .... \geq \sigma_r$, provides these basis vectors.

For a $k \leq m$ the POD-basis is defined as the set of the first k left singular vectors $\varphi_1, ..., \varphi_k$ of U. Let $V = [\varphi_1, ..., \varphi_k] \in \mathbb{R}^{n\times k}$ the matrix of these POD-basis vectors, a approximation of a snapshot y_j in the span of Y is therefore Vc with coefficients $c \in \mathbb{R}^k$. The Galerkin orthogonality of the residual $y_j - Vc$ to the span of V gives $V^T(y_j - Vc) = 0$, thereby the approximation $y_j \approx VV^Ty_j$ follows. The POD-basis provides a optimal orthogonal basis with regard to the sum of the quadratic error of the approximation of the snapshots.

Theorm 1 ***(POD-error bound):*** *Let* $\mathbb{V}_k = \{W \in \mathbb{R}^{n\times k} : W^TW = I_k\}$ *be the set of all k-dimensional orthonormal bases and* $V \in \mathbb{R}^{n\times k}$ *the matrix of basis vectors provided by the POD-method related to the snapshots* $Y = \{y_1, ..., y_m\}$*, then the following holds,*

$$\sum_{j=1}^{m} \parallel y_j - VV^Ty_j \parallel_2^2 = \min_{W\in\mathbb{V}_k} \sum_{j=1}^{m} \parallel y_j - WW^Ty_j \parallel_2^2 = \sum_{i=k+1}^{p} \sigma_i^2. \tag{2.8}$$

Theorm 2 ***(Schmidt-Eckart-Young):*** *Let* $A \in \mathbb{R}^{m\times n}$ *be a matrix of rank p, then there exist two orthogonal matrices* $U = [\varphi_1|...|\varphi_n] \in \mathbb{R}^{n\times n}$ *and* $Z = [\Psi_1|...|\Psi_m] \in \mathbb{R}^{m\times m}$*, so that* $A = U\Sigma Z^T$ *with the diagonal-matrix* $\Sigma = diag(\sigma_1, ..., \sigma_p)$ *and the singular values* $\sigma_1 \geq .... \geq \sigma_p \geq 0$*. The matrix*

$$A_k = \sum_{i=1}^{k} \sigma_i\varphi_i\Psi_i^T, \qquad 0 \leq k \leq p \tag{2.9}$$

satisfies the property

$$\|A - A_k\|_F = \min_{\substack{B\in\mathbb{R}^{m\times n} \\ rank(B)\leq k}} \|A - B\|_F = \sqrt{\sum_{i=k+1}^{p} \sigma_i^2}. \tag{2.10}$$

Proof of Theorem 1: By Theorem 2, the best rank k approximation of Y is given by

$$Y_k = \sum_{i=1}^{k} \sigma_i \varphi_i \Psi_i^T.$$

With $\Psi_i = \frac{1}{\sigma_i} Y^T \varphi_i$, which comes out of the definition of the SVD, follows

$$Y_k = \sum_{i=1}^{k} \sigma_i \varphi_i (\frac{1}{\sigma_i} Y^T \varphi_i)^T = \sum_{i=1}^{k} \varphi_i \varphi_i^T Y = VV^T Y.$$

Moreover, for a $W \in V_k$

$$\sum_{j=1}^{m} \| y_j - WW^T y_j \|_2^2 = \| Y - WW^T Y \|_F^2,$$

because of the definition of the Frobenius norm. And with (2.10) it is

$$\|Y - Y_k\|_F^2 = \min_{\substack{B\in\mathbb{R}^{m\times n} \\ rank(B)\leq k}} \|Y - B\|_F^2 \leq \min_{W\in\mathbb{V}_k} \| Y - WW^T Y \|_F^2 .$$

As $Y_k = VV^T Y$ follows

$$\| Y - VV^T Y \|_F^2 = \min_{W\in\mathbb{V}_k} \| Y - WW^T Y \|_F^2 = \sum_{i=k+1}^{p} \sigma_i^2.$$

The relative information of the first k POD-basis vectors can be described by

$$I(k) = \frac{\sum_{i=1}^{k} \sigma_i^2}{\sum_{i=1}^{p} \sigma_i^2} \tag{2.11}$$

and can be used as a error condition for choosing a k for a given approximation error ϵ_{POD}

$$I(k) = \frac{\sum_{i=1}^{k} \sigma_i^2}{\sum_{i=1}^{p} \sigma_i^2} \leq 1 - \epsilon_{POD}^2. \tag{2.12}$$

2.3 Discrete Empirical Interpolation Method (DEIM)

The projection of a nonlinear function $q(t) : \mathbb{D} \longmapsto \mathbb{R}^n$ onto a low-dimensional subspace spanned by a basis $U \in \mathbb{R}$ is of the form

$$q(t) \approx Uc(t), \tag{2.13}$$

where $c(t)$ is the corresponding coefficient vector. To determine the vector $c(t)$ there are m distinguished rows selected from the over-determined system $q(t) = Uc(t)$. The Selection of the rows can be done with a matrix

$$P = [e_{p1}, ..., e_{pm}] \in \mathbb{R}^{n \times m}, \tag{2.14}$$

where $e_{pi} = [0, ...0, 1, 0, ..., 0]^T \in \mathbb{R}^n$ is the pi-th unit vector, for $i = 1, .., m$. By multiplying P^T on $q(t) = Uc(t)$, m rows are selected and when $P^T U$ is non-singular, the vector $c(t)$ can be determined from

$$c(t) = (P^T U)^{-1} P^T q(t) \tag{2.15}$$

and the approximation 2.13 becomes

$$q(t) \approx Uc(t) = U(P^T U)^{-1} P^T q(t) \tag{2.16}$$

DEIM thus requires to constructs a projection basis U and interpolation indices $\{p_1, ..., p_m\}$ used for P. The projection basis U is constructed by using the POD on a nonlinear set of snapshots $Y_{deim} = [q(t_1), ..., q(t_m)]$ obtain from the high order system. These snapshots are the evaluations of the nonlinear function $q(t)$ on $t_1, ..., t_{n_s}$ points They can be obtained together with the snapshots for the full-system POD, because they are already computed there. So only the SVD has to be computed to obtain U.

The interpolation indices are iteratively selected by a greedy algorithm 1 from the basis U. The algorithm select iterative m indices $\{p_1, ..., p_m\}$, where $p_i \in \{1, ..., n\}$ for $i = 1, .., m$, which minimize the interpolation error over then snapshots set measured in the maximum norm. It starts by selecting the index of the biggest entry of the first column of U and thus the first basis vector, which is corresponding to the most dominant singular value of the SVD. Further indices are selected so that they correlate to the index of the larges entry of the residual $r = u_i - Uc$. The residual r is the error between the basis vector u_i and its projected approximation Uc. Since the columns of U are linear independent, r is in each step a nonzero vector. So an entry with the larges magnitude can always be selected.

Algorithm 1 DEIM

1: **Input:** $\{u_i\}_{i=1}^m \subset \mathbb{R}^n$
2: **Output:** $\vec{\varrho} = \{\varrho_1, .., \varrho_m\}^T \in \mathbb{N}^m$
3: $\varrho_1 = argmax\{|u_1|\}$
4: $U = [u_1], P = [e_{\varrho_1}], \vec{\varrho} = [\varrho_1]$
5: **for** $i = 2$ to m **do**
6:     **Solve** $(P^T U)c = P^T u_i$
7:     $r = u_i - Uc$
8:     $\varrho_i = argmax\{|r|\}$
9:     $U \leftarrow [U|u_i], P \leftarrow [P|e_{\varrho_i}], \vec{\varrho} \leftarrow \begin{bmatrix} \varrho \\ \varrho_i \end{bmatrix}$
10: **end for**

In line 6 in algorithm 1 there is to solve a linear system $(P^T U)c = P^T u_i$, which is only possible, if $P^T U$ is always non-singular, this is shown in [1]. The whole procedure has complexity $O(m^4 + mn)$, but since the dimension m of the basis U is small, it is not a problem. In chapter 3.4 the complexity of the POD-DEIM-method will be further discussed.

Formally the DEIM approximation is defined as follows.

Definition 2 *Let $f : D \longmapsto \mathbb{R}^n$ be a nonlinear vector valued function with $D \subset \mathbb{R}^d$ for a $d \in \mathbb{N}$. Let $\{u_l\}_{l=1}^m \subset \mathbb{R}^n$ be a set of linear independent vectors. For a $x \in D$ the DEIM approximation of order m in space spanned by $\{u_l\}_{l=1}^m$ is given by:*

$$\hat{f}(x) := U(P^T U)^{-1} P^T f(x) \tag{2.17}$$

where $U = [u_1|...|u_m] \in \mathbb{R}^{n\times m}$ is the basis, that the POD-method provides out of the nonlinear snapshots and $P = [e_{p_1}|...|e_{p_m}] \in \mathbb{R}^{n\times m}$, where $\{p_1, .., p_m\}$ being the output of the DEIM-algorithm with input of $\{u_l\}_{l=1}^m$.

Its can be shown, that $\hat{f}(x)$ is a interpolation of the original $f(x)$, by

$$P^T\hat{f}(x) = P^T(U(P^T U)^{-1}P^T f(x)) = (P^T U)(P^T U)^{-1}P^T f(x) = P^T f(x).$$

The functions $\hat{f}(x)$ and $f(x)$ are exactly the same at the DEIM points $\{p_1, .., p_m\}$.

Theorm 3 ***(DEIM-error):*** *Let $f \in \mathbb{R}^n$ be a vector, let $\{u_l\}_{l=1}^m \subset \mathbb{R}^n$ be a set of linear independent orthogonal vectors. The DEIM approximation of order $m \leq n$ of f in the space spanned by $\{u_l\}_{l=1}^m$, is $\hat{f} = \mathbb{P}f$, where $\mathbb{P} = U(P^T U)^{-1}P^T$, $U = [u_1|...|u_m] \in \mathbb{R}^{n\times m}$ and $P = [e_{p_1}|...|e_{p_m}] \in \mathbb{R}^{n\times m}$, with $\{p_1, .., p_m\}$ being the output of the DEIM-algorithm with input $\{u_l\}_{l=1}^m$. Then the following holds*

$$\| f - \hat{f} \|_2 \leq C_m \varepsilon(f) \tag{2.18}$$

where $C_m = \| (P^T U)^{-1} \|_2$ and $\varepsilon(f) = \| f - UU^T f \|_2$.

Proof of theorem 3: Let $\hat{f}(x)$ be the DEIM approximation, to find an error bound in the 2-norm for $\| f - \hat{f} \|_2$, look at the best approximation of f in the space spanned by U, in particular $f_* = UU^T f$. Then follows

$$f = (f - f_*) + f_* = w + f_* \qquad \text{with} \qquad w = f - f_* = f - UU^T f \tag{2.19}$$

and also

$$\hat{f} = \mathbb{P}f = \mathbb{P}(w + f_*) = \mathbb{P}w + f_*, \tag{2.20}$$

because $\mathbb{P}f_* = f_*$. Therefor follows

$$\| f - \hat{f} \|_2 = \| w + f_* - (\mathbb{P}w + f_*) \|_2 \leq \| I - \mathbb{P} \|_2 \| w \|_2 \,. \tag{2.21}$$

With $\| I - \mathbb{P} \|_2 = \| \mathbb{P} \|_2$, as shown in [2] for any projector-matrix $P \neq 0$, from 2.21 follows

$$\| f - \hat{f} \|_2 \leq \| I - \mathbb{P} \|_2 \| w \|_2 = \| \mathbb{P} \|_2 \| w \|_2 \tag{2.22}$$

and with $\| \mathbb{P} \|_2 = \| U(P^T U)^{-1} P^T \|_2 = \| (P^T U)^{-1} \|_2$, because U and $\mathbb{P}$ are both orthogonal, it follows

$$\| f - \hat{f} \|_2 \leq \| (P^T U)^{-1} \|_2 \| w \|_2 \tag{2.23}$$

Overall holds

$$\| f - \hat{f} \|_2 \leq \| \mathbb{P} \|_2 \| w \|_2 = \| (P^T U)^{-1} \|_2 \| f - UU^T f \|_2 = C_m \varepsilon(f). \tag{2.24}$$

The factor $(P^T U)^{-1}$ is dependent on the matrix P, which is constructed out of the DEIM indices $\{p_1, .., p_m\}$, thereby it explains that the DEIM algorithm aims to select an index to limit the growth of $\| (P^T U)^{-1} \|_2$ and so the error of $\| f - \hat{f} \|_2$. Chaturantabut shows in [1] a recursive formula for $(P^T U)^{-1}$, which proofs the minimization of this error and also the non-singularity of $(P^T U)^{-1}$, which is needed in the algorithm 1. Out of this also follows that C_m can be bound with

$$C_m \leq \frac{(1 + \sqrt{2n})^{m-1}}{\| u_1 \|_\infty}. \tag{2.25}$$

But this bound is very pessimistic and it will grow more rapidly than the actual matrix $(P^T U)^{-1}$, so it is not optimal for a priori error estimate, for a posteriori error estimate the computation of the norm of $(P^T U)^{-1}$ can be used, because the matrix is usually small.

The term $\varepsilon(f) = \| f - UU^T f \|_2$ is depended on $f(x)$ so it changes for ever new x, hence is expensive to compute. It would be desirable to have an easily computable estimation of it. A reasonable estimation is possible out of the SVD of the nonlinear snapshots

$$V = [q_1, ..., q_m], \qquad \text{where } q_i = q(y_j), i = 1, .., m \text{ and } j \in [0, 1). \tag{2.26}$$

Let $V = U\Sigma W^T$ be the SVD, where $U = [\hat{U}, \tilde{U}]$ with $\hat{U}$ representing the leading $k < m$ columns of U and $\tilde{U}$ representing the last $m - k$ columns of U and $W = [\hat{W}, \tilde{W}]$ equivalent to U. Partitioning $\Sigma = \begin{bmatrix} \hat{\Sigma} & 0 \\ 0 & \tilde{\Sigma} \end{bmatrix}$ that it corespont to the Partition of U and W. A vector $y \in Range(V)$ can be written as $y = Vc = \hat{U}\hat{\Sigma}\hat{W}^T c + \tilde{U}\tilde{\Sigma}\tilde{W}^T c$, with $c \in \mathbb{R}^m$. And it follows that

$$\begin{aligned}
\| y - y_* \| &= \| (I - \hat{U}\hat{U}^T) y \| \\
&= \| (I - \hat{U}\hat{U}^T)(\hat{U}\hat{\Sigma}\hat{W}^T c + \tilde{U}\tilde{\Sigma}\tilde{W}^T c) \| \\
&= \| \hat{U}\hat{\Sigma}\hat{W}^T c + \tilde{U}\tilde{\Sigma}\tilde{W}^T c - \hat{U}\underbrace{\hat{U}^T\hat{U}}_{=I}\hat{\Sigma}\hat{W}^T c - \hat{U}\underbrace{\hat{U}^T\tilde{U}}_{=0}\tilde{\Sigma}\tilde{W}^T c \| \\
&= \| \tilde{U}\tilde{\Sigma}\tilde{W}^T c \| \\
&\leq \sigma_{k+1} \| c \| .
\end{aligned}$$

For a vector y out of the $Range(V)$ the approximation is $y = Vc + w$ for a $w \in \mathbb{R}^n$ and $w^T Vc = 0$, if $\| w \|$ is in $\mathcal{O}(\sigma_{k+1})$, then $\varepsilon(y) \leq \sigma_{k+1} \| c \| + 2 \| w \| \approx \sigma_{k+1}$. The DEIM Error is bound like this, if the trajectories are attrected to a low-dimensional subspace, this means they lie nearly in the space spanned by the snapshots.