POD and DEIM in field-flow fractionation

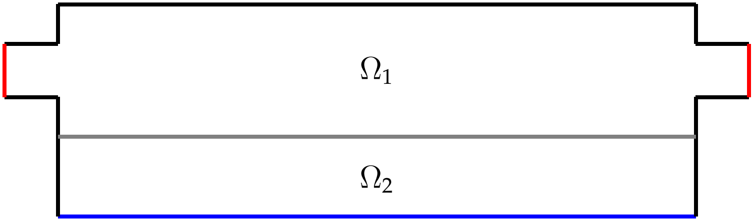
Carina Willbold
Universität Augsburg

Sion
August 1, 2012

Outline

Stokes–Brinkman equation

$$
\begin{aligned}
\rho \frac{\partial \mathbf{v}}{\partial t} - \nu \Delta \mathbf{v} + \nu \chi_{\Omega_2} K^{-1} \mathbf{v} + \nabla p &= \mathbf{0} && \text{in} \quad \Omega \times (0,T) \\
\nabla \cdot \mathbf{v} &= 0 && \text{in} \quad \Omega \times (0,T) \\
\mathbf{v} &= \mathbf{v}_{\text{in}}^{(i)} && \text{on} \quad \Gamma_{\text{in}}^{(i)} \times (0,T), 1 \le i \le 2 \\
\mathbf{v} &= \mathbf{0} && \text{on} \quad \Gamma_{\text{lat}} \times (0,T) \\
\nu \frac{\partial \mathbf{v}}{\partial \mathbf{n}_\Gamma} - p \mathbf{n}_\Gamma &= \mathbf{0} && \text{on} \quad \Gamma_{\text{out}} \times (0,T) \\
\mathbf{v}(\cdot, 0) &= \mathbf{v}_0 && \text{in} \quad \Omega
\end{aligned}
$$



Stokes–Brinkman equation

weak formulation

Find $(\mathbf{v}, p) \in W(0,T) \times Q(0,T)$ such that

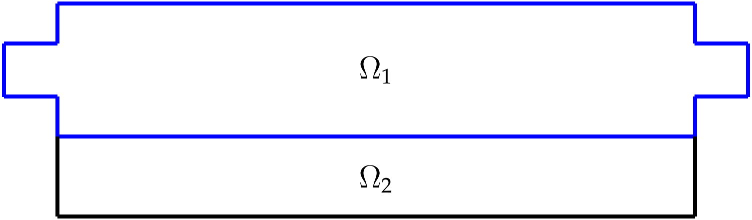
$$\frac{\partial}{\partial t}(\mathbf{v}, \mathbf{z})_{0,\Omega} + a_s(\mathbf{v}, \mathbf{z}) - b_s(p, \mathbf{z}) = 0, \quad t \in (0,T]$$
$$b_s(z, \mathbf{v}) = 0, \quad t \in (0,T]$$
$$(\mathbf{v}(\cdot,0), \mathbf{z})_{0,\Omega} = (\mathbf{v}^0, \mathbf{z})_{0,\Omega}$$

holds true for all $\mathbf{z} \in V_0$, $z \in Q(0,T)$, where

$$V := \{\mathbf{v} \in H^1(\Omega) \mid \mathbf{v}|_{\Gamma_{\mathrm{in}}^{(i)}} = \mathbf{v}_{\mathrm{in}}^{(i)}, \mathbf{v}|_{\Gamma_{\mathrm{lat}}} = \mathbf{0}\}$$
$$V_0 := \{\mathbf{v} \in H^1(\Omega) \mid \mathbf{v}|_{\Omega \setminus \Gamma_{\mathrm{out}}} = \mathbf{0}\},$$
$$W(0,T) := H^1(0,T;V^*) \cap L^2(0,T;V),$$
$$Q(0,T) := L^2(0,T;L^2(\Omega)).$$

Advection diffusion equation

$$
\begin{aligned}
\frac{\partial c}{\partial t} - \nabla \cdot D \nabla c + \mathbf{v} \cdot \nabla c &= 0 && \text{in} \quad \Omega_1 \times (0,T) \\
\frac{\partial c}{\partial \mathbf{n}_\Gamma} &= 0 && \text{on} \quad \Gamma_1 \times (0,T) \\
c(\cdot, 0) &= c_0 && \text{in} \quad \Omega_1
\end{aligned}
$$



Advection diffusion equation

weak formulation

Find $c \in W(0,T)$ such that

$$\frac{\partial}{\partial t}(c,z)_{0,\Omega_1} + a_d(c,z) = 0, \quad t \in (0,T]$$
$$(c(\cdot,0),z)_{0,\Omega_1} = (c^0,z)_{0,\Omega_1}$$

holds true for all $z \in H^1(\Omega_1)$.

Advection diffusion equation

discretization in space

Consider

$$V_h := \{v_h \in C(\Omega_1) \mid v_h|_K \in P_2(K), K \in \mathcal{T}_h(\Omega_1)\}.$$

Find then $\mathbf{c} \in C^1(0,T;V_h)$ s. t. it holds for all $\psi \in V_h$:

$$\sum_{K\in\mathcal{T}_h(\Omega_1)} (\frac{\mathrm{d}\mathbf{c}}{\mathrm{d}t}, \psi + \tau_K h_K(\mathbf{v}\cdot\nabla\psi))_{0,K} + (D\nabla\mathbf{c}, \nabla\psi)_{0,\Omega_1} +$$

$$(\mathbf{v}\cdot\nabla\mathbf{c}, \psi)_{0,\Omega_1} +$$

$$\sum_{K\in\mathcal{T}_h(\Omega_1)} \tau_K h_K(-D\Delta\mathbf{c} + \mathbf{v}\cdot\nabla\mathbf{c}, \mathbf{v}\cdot\nabla\psi)_{0,K} = 0$$

$$(\mathbf{c}(\cdot,0), \psi)_{0,\Omega_1} = (\mathbf{c}^0, \psi)_{0,\Omega_1}.$$

Discretization in space

$$\min \quad J(\mathbf{u}) = \frac{1}{2}\|c(\cdot,T) - c^{\text{foc}}\|^2_{0,\Omega_1}$$

$\downarrow$ $\mathbf{u}$

$$\begin{aligned} M_s \frac{\partial \mathbf{v}_h(t)}{\partial t} + A_s \mathbf{v}_h(t) - B_s^T p_h(t) + C_s \mathbf{u}(t) &= 0, \quad && t \in (0,T] \\ B_s \mathbf{v}_h(t) &= 0, \quad && t \in (0,T] \\ M_s \mathbf{v}_h(0) &= \mathbf{v}_h^0 && \end{aligned}$$

$\downarrow$ $\mathbf{v}_h$

$$\begin{aligned} (M_d^1 + M_d^2(\mathbf{v}_h)) \frac{\partial c_h(t)}{\partial t} + A_d(\mathbf{v}_h) c_h(t) &= 0, \quad && t \in (0,T] \\ M_d^1 c_h(0) &= c_h^0 && \end{aligned}$$

Motivation

We want to solve

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}(t) = A\mathbf{y}(t) + F(\mathbf{y}(t))$$
$$\mathbf{y}(0) = \mathbf{y}_0$$

with varying values of parameters, where $\mathbf{y}(t) \in \mathbb{R}^n$ is unknown and $A \in \mathbb{R}^{n\times n}$, $F\colon \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{y}_0 \in \mathbb{R}^n$ are given.

Such problems arise in iterative algorithms for solving optimization problems with differential equations as constraints.

Motivation

We want to solve

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{y}(t) = A\mathbf{y}(t) + F(\mathbf{y}(t))$$
$$\mathbf{y}(0) = \mathbf{y}_0$$

with varying values of parameters, where $\mathbf{y}(t) \in \mathbb{R}^n$ is unknown and $A \in \mathbb{R}^{n\times n}$, $F\colon \mathbb{R}^n \to \mathbb{R}^n$, $\mathbf{y}_0 \in \mathbb{R}^n$ are given.

Empirical observation: The solution trajectories are often approximately contained in a low-dimensional subspace of $\mathbb{R}^n$.

Plan

Main idea: Because we do not need the full flexibility of $\mathbb{R}^n$, we can instead search for the solutions in a subspace having much smaller dimension $\ell \ll n$.

Snapshots

Question:

How to find the low-dimensional supspace of $\mathbb{R}^n$ which approximately contains the trajectories?

Snapshots

Question:
How to find the low-dimensional supspace of $\mathbb{R}^n$ which approximately contains the trajectories?

One possibility:
Record snapshots $\mathbf{y}_i := \mathbf{y}(t_i) \in \mathbb{R}^n$ of a "typical" trajectory at certain times $t_1, \ldots, t_{n_s}$. Then find an ℓ-dimensional subspace with smallest distance to the snapshots.

Proper Orthogonal Decomposition

Find an ℓ-dimensional subspace, $\ell \ll n$, with the smallest distance to the snapshots:

Minimize

$$\sum_{i=1}^{n_s} \|\mathbf{y}_i - P_V(\mathbf{y}_i)\|^2$$

such that

$$V \in \text{set of all } \ell\text{-dimensional subspaces of } \mathbb{R}^n,$$

where P_V denotes the orthogonal projection onto V.

Proper Orthogonal Decomposition

Find an ℓ-dimensional subspace, $\ell \ll n$, with the smallest distance to the snapshots:

Better formulation: Minimize

$$J(\mathbf{v}_1, \dots, \mathbf{v}_\ell) := \sum_{i=1}^{n_s} \left\| \mathbf{y}_i - \sum_{k=1}^{\ell} \langle \mathbf{y}_i, \mathbf{v}_k \rangle \mathbf{v}_k \right\|^2$$

such that

$$\mathbf{v}_1, \dots, \mathbf{v}_\ell \in \mathbb{R}^n$$
$$\langle \mathbf{v}_j, \mathbf{v}_k \rangle = \delta_{jk}, \ 1 \le j, k \le \ell.$$

Numerical solution

1. by using the singular value decomposition of $Y := (\mathbf{y}_1 | \cdots | \mathbf{y}_{n_s}) \in \mathbb{R}^{n \times n_s}$
2. by using the eigenvalue decomposition of $\mathcal{K} := ((\mathbf{y}_i, \mathbf{y}_j))_{i,j} \in \mathbb{R}^{n_s \times n_s}$

Solution by singular value decomposition

Theorem (Singular value decomposition, ca. 1873)

Let $Y \in \mathbb{R}^{n \times n_s}$. Then there exist orthogonal matrices $\widetilde{V} \in \mathbb{R}^{n \times n}$ and $W \in \mathbb{R}^{n_s \times n_s}$ with

$$Y = \widetilde{V} \Sigma W^T,$$

where $\Sigma \in \mathbb{R}^{n \times n_s}$ is a (rectangular) diagonal matrix. The diagonal elements are called the singular values. They are nonnegative and given in descending order.

Then a solution of the optimization problem is given by the first ℓ columns of $\widetilde{V}$, when $Y := (\mathbf{y}_1 | \cdots | \mathbf{y}_{n_s}) \in \mathbb{R}^{n \times n_s}$ denotes the matrix of the snapshots.

Solution by eigenvalue decomposition

Consider the eigenvalue decomposition of $\mathcal{K}$:

$$\mathcal{K} = TDT^* \in \mathbb{R}^{n_s \times n_s}.$$

Then the POD basis vectors are given by

$$\mathbf{u}_i := \frac{1}{\sqrt{D_{ii}}} Y \, \mathbf{t}_i$$

for $1 \leq i \leq N$, $D_{ii} \neq 0$, with $T = (\mathbf{t}_1 | \cdots | \mathbf{t}_{n_s}) \in \mathbb{R}^{n_s \times n_s}$ and the snapshot matrix $Y := (\mathbf{y}_1 | \cdots | \mathbf{y}_{n_s})$.

Approximation error

The approximation error is given by

$$J(\mathbf{v}_1, \ldots, \mathbf{v}_k) = \sum_{i=1}^{n_s} \|\mathbf{y}_i\|^2 - \sum_{k=1}^{\ell} \sigma_k^2,$$

where σ_j denotes the singular values of $Y = (\mathbf{y}_1 | \cdots | \mathbf{y}_{n_s}) \in \mathbb{R}^{n \times n_s}$.

Thus, the POD basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is ordered by importance.

Model reduction by projection

Let $\mathbf{v}_1, \ldots, \mathbf{v}_\ell \in \mathbb{R}^n$ denote an orthonormal basis of a low-dimensional subspace of $\mathbb{R}^n$, $\ell \ll n$, and let $V_\ell := (\mathbf{v}_1 | \cdots | \mathbf{v}_\ell) \in \mathbb{R}^{n \times \ell}$. We make the ansatz

$$\mathbf{y}(t) := V_\ell \tilde{\mathbf{y}}(t),$$

with $\tilde{\mathbf{y}} \in \mathbb{R}^\ell$. The Galerkin projection is then given by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \tilde{\mathbf{y}}(t) &= V_\ell^T A V_\ell \tilde{\mathbf{y}}(t) + V_\ell^T F(V_\ell \tilde{\mathbf{y}}(t)), \\ \tilde{\mathbf{y}}(0) &= V_\ell^T \mathbf{y}_0, \end{aligned}$$

in which the linear part is reduced ($V_\ell^T A V_\ell \in \mathbb{R}^{\ell \times \ell}$), but the nonlinear one is not.

Discrete Empirical Interpolation Method

[Chaturantabut, Sorensen (2009)]

The partially reduced order model is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathbf{y}}(t) = V_\ell^T A V_\ell \widetilde{\mathbf{y}}(t) + V_\ell^T F(V_\ell \widetilde{\mathbf{y}}(t)),$$
$$\widetilde{\mathbf{y}}(0) = V_\ell^T \mathbf{y}_0.$$

For example let $F\colon \mathbb{R}^n \to \mathbb{R}^n$ have the form

$$F\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix},$$

where $f\colon \mathbb{R} \to \mathbb{R}$.

Discrete Empirical Interpolation Method

[Chaturantabut, Sorensen (2009)]

For example let $F\colon \mathbb{R}^n \to \mathbb{R}^n$ have the form

$$F\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} f(x_1) \\ \vdots \\ f(x_n) \end{pmatrix},$$

where $f\colon \mathbb{R} \to \mathbb{R}$.

To construct the reduced order model, we record snapshots $\mathbf{z}_i = F(\mathbf{y}(t_i)) \in \mathbb{R}^n$, $1 \leq i \leq n_s$ and determine a POD basis $\mathbf{u}_1, \ldots, \mathbf{u}_k \in \mathbb{R}^n$, $k \ll n$.

Projection of the nonlinearity

Let $\mathbf{u}_1, \ldots, \mathbf{u}_k \in \mathbb{R}^n$, $k \ll n$ be a POD-basis and set $U_k := (\mathbf{u}_1 | \cdots | \mathbf{u}_k) \in \mathbb{R}^{n \times k}$.

Recall the partially reduced order model

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathbf{y}}(t) = V_\ell^T A V_\ell \widetilde{\mathbf{y}}(t) + V_\ell^T F(V_\ell \widetilde{\mathbf{y}}(t)).$$

First idea: The equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathbf{y}}(t) = V_\ell^T A V_\ell \widetilde{\mathbf{y}}(t) + V_\ell^T U_k U_k^T F(V_\ell \widetilde{\mathbf{y}}(t))$$

should still be a good approximation, because $U_k U_k^T$ is the orthogonal projection onto the span of $\mathbf{u}_1, \ldots, \mathbf{u}_k$.

Projection of the nonlinearity

We search for a projection matrix $Q \in \mathbb{R}^{n \times n}$ onto the span of the optimal basis $\mathbf{u}_1, \dots, \mathbf{u}_k$, such that the nonlinearity $V_\ell^T Q F(V_\ell \widetilde{\mathbf{y}}(t))$ can be evaluated efficiently. To define $Q\mathbf{x}$ for $\mathbf{x} \in \mathbb{R}^n$, we make the ansatz

$$U_k \mathbf{c} \approx \mathbf{x}$$

where the coefficient vector $\mathbf{c} \in \mathbb{R}^k$ is unknwon. We can only approximately fulfill a certain choice $p_1, \dots, p_k$ of the equations:

$$P^T U_k \mathbf{c} \approx P^T \mathbf{x}$$

with $P := (\mathbf{e}_{p_1} | \cdots | \mathbf{e}_{p_k}) \in \mathbb{R}^{n \times k}$ and $U = (\mathbf{u}_1 | \cdots | \mathbf{u}_\ell) \in \mathbb{R}^{n \times k}$.

Then we set $Q\mathbf{x} := U_k \mathbf{c} = U_k (P^T U_k)^{-1} P^T \mathbf{x} \in \mathbb{R}^n$.

Reduced order model

Recall the partially reduced order model

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathbf{y}}(t) = V_\ell^T A V_\ell \widetilde{\mathbf{y}}(t) + V_\ell^T F(V_\ell \widetilde{\mathbf{y}}(t)),$$
$$\widetilde{\mathbf{y}}(0) = V_\ell^T \mathbf{y}_0.$$

With the projection matrix

$$Q = U_k (P^T U_k)^{-1} P^T \in \mathbb{R}^{n \times n}$$

we obtain the fully reduced order model

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{\mathbf{y}}(t) = V_\ell^T A V_\ell \widetilde{\mathbf{y}}(t) + V_\ell^T U_k (P^T U_k)^{-1} P^T F(V_\ell \widetilde{\mathbf{y}}(t))$$
$$\widetilde{\mathbf{y}}(0) = V_\ell^T \mathbf{y}_0.$$

Selection algorithm

Input: orthonormal basis $\mathbf{u}_1, \ldots, \mathbf{u}_k \in \mathbb{R}^n$

Output: selection matrix P such that $U_k(P^T U_k)^{-1} P^T$ is a good projection matrix

1. $\mathbf{r}_1 := \mathbf{u}_1$.
   $p_1 := \arg\min_{i=1,\ldots,n} |(\mathbf{r}_1)_i|$, $P := (\mathbf{e}_{p_1}) \in \mathbb{R}^{n\times 1}$.
2. $\mathbf{r}_2 := \mathbf{u}_2 - \widetilde{U}(P^T\widetilde{U})^{-1}P^T\mathbf{u}_2$ with $\widetilde{U} = (\mathbf{u}_1) \in \mathbb{R}^{n\times 1}$.
   $p_2 := \arg\min_{i=1,\ldots,n} |(\mathbf{r}_2)_i|$, $P := (\mathbf{e}_{p_1}|\mathbf{e}_{p_2}) \in \mathbb{R}^{n\times 2}$.
3. $\mathbf{r}_3 := \mathbf{u}_3 - \widetilde{U}(P^T\widetilde{U})^{-1}P^T\mathbf{u}_3$ with $\widetilde{U} = (\mathbf{u}_1|\mathbf{u}_2) \in \mathbb{R}^{n\times 2}$.
   $p_3 := \arg\min_{i=1,\ldots,n} |(\mathbf{r}_3)_i|$, $P := (\mathbf{e}_{p_1}|\mathbf{e}_{p_2}|\mathbf{e}_{p_3}) \in \mathbb{R}^{n\times 3}$.
4. and so on ...

Error estimate

Theorem (Chaturantabut, Sorensen (2009))

For the approximation

$$\widehat{F}(\mathbf{x}) := U_k(P^T U_k)^{-1} P^T F(\mathbf{x})$$

holds the error estimate

$$\|F(\mathbf{x}) - \widehat{F}(\mathbf{x})\|_2 \leq \|(P^T U_k)^{-1}\|_2 \, \|(I - U_k U_k^T)\mathbf{F}(x)\|_2.$$

Up to the constant $\|(P^T U_k)^{-1}\|_2$ we have just the consistency error

$$\|(I - U_k U_k^T)\mathbf{F}(x)\|_2 = \min_{\mathbf{v} \in \operatorname{im} U_k} \|\mathbf{F}(x) - \mathbf{v}\|_2.$$

Error estimate

The DEIM algorithm garanties that $P^T U_k$ is invertible. The term $\|(P^T U_k)^{-1}\|_2$ can be estimated by a constant, but the bound is very conservative.

DEIM in a general Hilbert space

We want to reduce a general nonlinear function

$$F\colon K \longrightarrow H$$

where K is an arbitrary parameter space and H is a finite-dimensional Hilbert space.

For example, we use $H := \mathbb{R}^{n_c \times n_c}$ for reducing the advection diffusion equation.

DEIM in a general Hilbert space

1. Record Snapshots $z_1, \dots, z_{n_s} \in H$.
2. Calculate a POD basis $u_1, \dots, u_k \in H$ by either
    1. using the operator singular value decomposition of

    $$\mathcal{R}\colon \begin{array}{ccc} H & \longrightarrow & H \\ x & \longmapsto & \sum_{i=1}^{n_s} (x, z_i)_H z_i \end{array}$$

    2. or using the eigenvalue decomposition of the matrix

    $$\mathcal{K} := ((z_i, z_j)_H)_{ij} \in \mathbb{R}^{n_s \times n_s}.$$

3. Apply the DEIM algorithm to $\mathcal{I}(u_1), \dots, \mathcal{I}(u_k)$ where $\mathcal{I}\colon H \to \mathbb{R}^n$ is a suitable isomorphism.

Error estimate

Then we obtain the error estimate

$$\|F(\mathbf{x}) - \widehat{F}(\mathbf{x})\|_H \leq \|\mathcal{I}^{-1}\| \, \|(P^T U_k)^{-1}\|_2 \, \|(I - U_k U_k^T)\mathcal{I}(\mathbf{F}(x))\|_2,$$

which depends on the isomorphism $\mathcal{I}$.