MAX PLANCK INSTITUTE
FOR DYNAMICS OF COMPLEX
TECHNICAL SYSTEMS
MAGDEBURG

CSC
COMPUTATIONAL METHODS IN
SYSTEMS AND CONTROL THEORY

Space-time Galerkin POD for optimal control of Burgers' equation

Manuel Baumann   Peter Benner   Jan Heiland

April 27, 2017

Absolventen Seminar • Numerische Mathematik, TU Berlin

Outline

Introduction

$$\dot{x} - \Delta x = f(t, x)$$

Consider the solution of a PDE:

$$x \in L^2(I; L^2(\Omega))$$

with $I \subset \mathbb{R}$ ... the time-interval
$\Omega \subset \mathbb{R}^n$ ... the spatial domain

and its numerical approximation:

$$\mathbf{x} \in \mathcal{S} \cdot \mathcal{Y}$$

with $\mathcal{S} \subset L^2(I)$ ... discretized time
$\mathcal{Y} \subset L^2(\Omega)$ ... a FE space

Task: Find $\hat{\mathcal{S}} \subset \mathcal{S}$ and $\hat{\mathcal{Y}} \subset \mathcal{Y}$ of much smaller dimension to express $\mathbf{x}$.

Space-Time Spaces

PDE solution $x \in L^2(I; L^2(\Omega))$
$\mathcal{S} \subset L^2(I)$ ... discretized time
$\mathcal{Y} \subset L^2(\Omega)$ ... a FE space

Consider finite dimensional subspaces

$$\mathcal{S} = \mathrm{span}\{\psi_1, \cdots, \psi_s\} \subset L^2(I)$$
$$\mathcal{Y} = \mathrm{span}\{\nu_1, \cdots, \nu_q\} \subset L^2(\Omega)$$

with the mass matrices

$$\mathbf{M}_{\mathcal{S}} = \left[(\psi_i, \psi_j)_{L^2}\right]_{i,j=1,\dots,s} \quad \text{and} \quad \mathbf{M}_{\mathcal{Y}} = \left[(\nu_i, \nu_j)_{L^2}\right]_{i,j=1,\dots,q}$$

and the product space

$$\mathcal{S} \cdot \mathcal{Y} \subset L^2(I; L^2(\Omega)).$$

Space-Time Spaces

We represent a function

$$\mathbf{x} = \sum_{j=1}^{s}\sum_{i=1}^{q} \mathbf{x}_{i\cdot j}\nu_i\psi_j \in \mathcal{S}\cdot\mathcal{Y}$$

via its matrix of coefficients

$$\mathbf{X} = \left[\mathbf{x}_{i\cdot j}\right]_{i=1,\dots,q}^{j=1,\dots,s} \in \mathbb{R}^{q,s}$$

and vice versa.

Optimal Bases

Lemma (Optimal low-rank bases in space[1])

Given $x \in \mathcal{S} \cdot \mathcal{Y}$ *and the associated matrix of coefficients* $\mathbf{X}$. *The best-approximating subspace* $\hat{\mathcal{Y}}$ *in the sense that* $\|\Pi_{\mathcal{S}\cdot\hat{\mathcal{Y}}}x - x\|_{\mathcal{S}\cdot\mathcal{Y}}$ *is minimal over all subspaces of* $\mathcal{Y}$ *of dimension* $\hat{q}$ *is given as* $\operatorname{span}\{\hat{\nu}_i\}_{i=1,\dots,\hat{q}}$, *where*

$$\begin{bmatrix} \hat{\nu}_1 \\ \hat{\nu}_2 \\ \vdots \\ \hat{\nu}_{\hat{q}} \end{bmatrix} = V_{\hat{q}}^{\mathsf{T}} \mathbf{M}_{\mathcal{Y}}^{-1/2} \begin{bmatrix} \nu_1 \\ \nu_2 \\ \vdots \\ \nu_q \end{bmatrix},$$

where $V_{\hat{q}}$ *is the matrix of the* $\hat{q}$ *leading left singular vectors of*

$$\mathbf{M}_{\mathcal{Y}}^{1/2} \mathbf{X} \mathbf{M}_{\mathcal{S}}^{1/2}.$$

[1] BM&PB&JH '16: *ArXiv:1611.04050*

Optimal Bases

The same arguments apply to the transpose of $\mathbf{X}$:

Lemma (Optimal low-rank bases in time[2])

Given $x \in \mathcal{S} \cdot \mathcal{Y}$ and the associated matrix of coefficients $\mathbf{X}$. The best-approximating subspace $\hat{\mathcal{S}}$ in the sense that $\|\Pi_{\hat{\mathcal{S}}\cdot\mathcal{Y}} x - x\|_{\mathcal{S}\cdot\mathcal{Y}}$ is minimal over all subspaces of $\mathcal{S}$ of dimension $\hat{s}$ is given as $\operatorname{span}\{\hat{\psi}_j\}_{j=1,\dots,\hat{s}}$, where

$$\begin{bmatrix} \hat{\psi}_1 \\ \hat{\psi}_2 \\ \vdots \\ \hat{\psi}_{\hat{s}} \end{bmatrix} = U_{\hat{s}}^{\mathsf{T}} \mathbf{M}_{\mathcal{S}}^{-1/2} \begin{bmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_s \end{bmatrix},$$

where $U_{\hat{s}}$ is the matrix of the $\hat{s}$ leading right singular vectors of

$$\mathbf{M}_{\mathcal{Y}}^{1/2} \mathbf{X} \mathbf{M}_{\mathcal{S}}^{1/2}.$$

[2] BM&PB&JH '16: *ArXiv:1611.04050*

Relation to POD

The solution of a spatially discretized PDE

$$x \colon \tau \mapsto \mathbb{R}^q$$

is projected to $\mathcal{S} \cdot \mathbb{R}^q$ via

$$\Pi_{\mathcal{S}\cdot\mathcal{Y}} x = \begin{bmatrix} (x_1, \psi_1)_{L^2} & \dots & (x_1, \psi_s)_{L^2} \\ \vdots & \ddots & \vdots \\ (x_q, \psi_1)_{L^2} & \dots & (x_q, \psi_s)_{L^2} \end{bmatrix} \mathbf{M}_{\mathcal{S}}^{-1}.$$

In the (degenerated) case that ψ_j is a delta distribution centered at $\tau_j \in I$, the coefficient matrix degenerates to

$$\begin{bmatrix} x_1(\tau_1) & \dots & x_1(\tau_s) \\ \vdots & \ddots & \vdots \\ x_q(\tau_1) & \dots & x_q(\tau_s) \end{bmatrix}$$

– the standard POD snapshot matrix.

Section 4

Space-Time Galerkin-POD for Optimal Control

Target 1: Step function

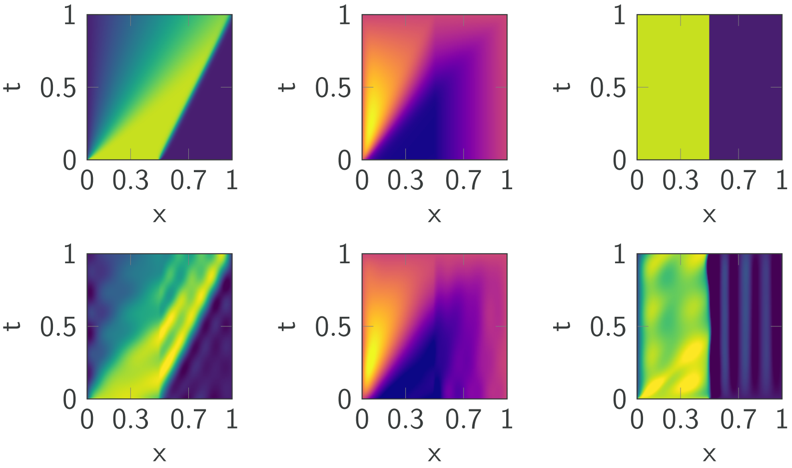

Figure : Illustration of the state, the adjoint, and the target and their approximation via POD-reduced space-time bases.

Target 2: Heart shape

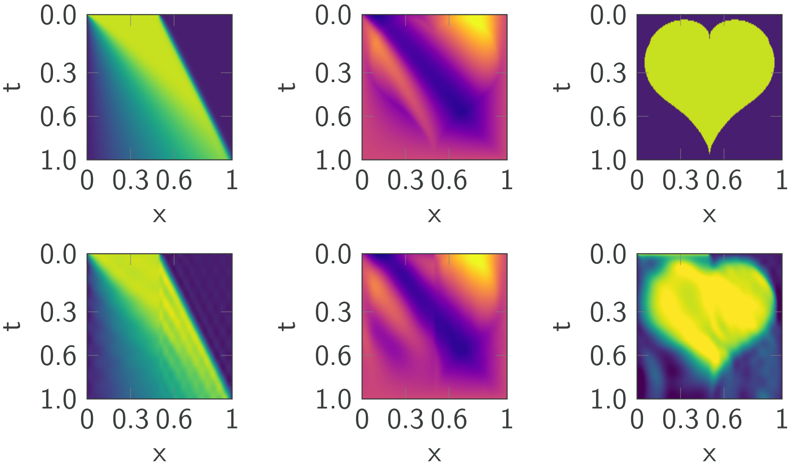

Figure : Illustration of the state, the adjoint, and the target and their approximation via POD-reduced space-time bases.

Space-Time Galerkin-POD for Optimal Control

For a target trajectory $x^* \in L^2(0, T; L^2(\Omega))$ and a penalization parameter $\alpha > 0$, consider

$$\mathcal{J}(x, u) := \frac{1}{2}\|x - x^*\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 \to \min_{u \in L^2(0,T;L^2(\Omega))}$$

subject to the generic PDE

$$\dot{x} - \Delta x + N(x) = f + u, \quad x(0) = 0. \tag{FWD}$$

If the nonlinearity is smooth, then necessary optimality conditions for (x, u) are given through $u = \frac{1}{\alpha}\lambda$, where λ solves the adjoint equation

$$-\dot{\lambda} - \Delta\lambda + D_x N(x)^{\mathsf{T}}\lambda + x = x^*, \quad \lambda(T) = 0. \tag{BWD}$$

Algorithm:

1. Do standard forward/backward solves to compute the matrix of measurements for x and λ.
2. Compute optimal low-dimensional spaces $\hat{\mathcal{S}}$, $\hat{\mathcal{R}}$, $\hat{\mathcal{Y}}$, and $\hat{\Lambda}$ for the space and time discretization of the state x and the adjoint state λ.
3. Solve the space-time Galerkin projected necessary optimality conditions (FWD)-(BWD)[3] for the reduced costate $\hat{\boldsymbol{\lambda}}$.
4. Define the suboptimal control via $\hat{\mathbf{u}} = \frac{1}{\alpha}\hat{\boldsymbol{\lambda}}$ and inflate it to the full space.
5. Apply it in the full order simulation.

[3](FWD)-(BWD) is a two-point boundary value problem with initial and terminal conditions for which time stepping schemes like RKM do not apply.

Numerical Setup

The PDE

- 1D Burger's equation
- $I = (0, 1]$, $\Omega = (0, 1)$
- Viscosity: $\nu = 5 \cdot 10^{-3}$
- Stepfunction as initial value
- Zero Dirichlet conditions

The optimization

- target 1: keep the initial state
- target 2: make a heart
- parameter: $\alpha = 10^{-3}$

The full model

- Equidistant space and time grids
- $\mathcal{S} = \mathcal{R}$ ... 120 linear hat functions
- $\mathcal{Y} = \Lambda$ ... 220 linear hat functions

The reduced model

- $\hat{\mathcal{Y}} = \hat{\Lambda}$ ... of dimension $\hat{q} = \hat{p}$
- $\hat{\mathcal{S}} \neq \hat{\mathcal{R}}$ ... of dimensions $\hat{s} = \hat{r}$
- $\hat{q}$, $\hat{p}$, $\hat{s}$, $\hat{r}$ ... varying
- n_t ... varying[4]

[4] dimension of time parametrization for an gradient based approach

Target 1: Step function

$\hat{K}$	24	36	48	72	96
$\frac{1}{2}\|\hat{x} - x_0\|_{L^2}^2$	0.0330	0.0280	0.0192	0.0121	0.0104
$\mathcal{J}(\hat{x}, \hat{u})$	0.0351	0.0309	0.0234	0.0177	0.0152
walltime $[s]$	0.1	0.48	1.81	18.7	155

Table : Performance of the suboptimal control versus the cumulative dimension $\hat{K} = \hat{p} + \hat{q} + \hat{r} + \hat{s}$ of the reduced bases with $\hat{p} = \hat{q} = \hat{r} = \hat{s}$.

$(\hat{q}, \hat{s})/(\hat{p}, \hat{r})$	(16, 7)	(15,10)	(12,10)	(10,12)	(10,15)	(7,16)
$\frac{1}{2}\|\hat{x} - x_0\|_{L^2}^2$	0.0143	0.0106	0.0151	0.0303	0.0318	0.0357
$\mathcal{J}(\hat{x}, \hat{u})$	0.0189	0.0159	0.0192	0.0340	0.0353	0.0382
walltime	1.43	2.71	1.58	1.42	2.54	1.11

Table : Performance of the suboptimal control versus varying distributions of space and time resolutions.

Target 1: Step function

$(\hat{q},\hat{s})/(\hat{p},\hat{r})$	(16, 7)	(15,10)	(12,10)	(10,12)	(10,15)	(7,16)
$\frac{1}{2}\|\hat{x}-x_0\|^2_{L^2}$	0.0143	0.0106	0.0151	0.0303	0.0318	0.0357
$\mathcal{J}(\hat{x},\hat{u})$	0.0189	0.0159	0.0192	0.0340	0.0353	0.0382
`walltime`	1.43	2.71	1.58	1.42	2.54	1.11

Table : Performance of the suboptimal control versus varying distributions of space and time resolutions.

$(\hat{q}, n_t)$	$(13,18)$	$(15,19)$	$(16,20)$	$(19,15)$	$(20,16)$	$(18,13)$
$\frac{1}{2}\|\hat{x}-x_0\|^2_{L^2}$	0.0144	0.0119	0.0110	0.0090	0.0091	0.0087
$\mathcal{J}(\hat{x},\hat{u})$	0.0146	0.0122	0.0113	0.0093	0.0095	0.0091
`walltime`	21.98	35.04	40.34	41.93	46.54	50.62

Table : Benchmark of an gradient based approach (SQP-POD-BFGS with $\alpha = 6.25 \cdot 10^{-5}$)

Conclusion

- The space-time Galerkin POD approach allows for
  - construction of optimized Galerkin bases in space and time
  - in a functional analytical framework
- The resulting space-time Galerkin discretization
  - approximates PDEs by a small system of algebraic equations
  - and naturally extends to boundary value problems in time
  - can be used for efficient computations of (sub)optimal controls
- Future work:
  - Use the functional analytical framework for error estimates.
  - Exploit the freedom of the choice of the measurement functions in $\mathcal{Y}$,
  - to produce, e.g., *optimal* measurements or to compensate for stochastic perturbations.

Further Reading and Coding

M. Baumann, P. Benner, and J. Heiland.
Space-Time Galerkin POD with application in optimal control of semi-linear parabolic partial differential equations.
ArXiv:1611.04050, Nov. 2016.

M. Baumann, J. Heiland, and M. Schmidt.
Discrete input/output maps and their relation to POD.
In P. Benner et al., editors, *Numerical Algebra, Matrix Theory, Differential-Algebraic Equations and Control Theory*, pages 585–608. Springer, 2015.

J. Heiland and M. Baumann.
spacetime-galerkin-pod-bfgs-tests – Python/Matlab implementation space-time POD and BFGS for optimal control of Burgers equation.
2016, `doi:10.5281/zenodo.166339`.

Thank you!

Thank you for your attention!

I am always open for discussion

heiland@mpi-magdeburg.mpg.de
www.janheiland.de
github.com/highlando