Optimality System Proper Orthogonal Decomposition for Optimal Control Problems with Control and State Constraints

Seminar on Optimization

Martin Gubisch

University of Konstanz

October 8, 2014

Universität Konstanz

Outline

Universität Konstanz

Problem formulation

We consider the optimal control problem

$$\min_{y,u,w} J(y,u,w) = \int_\Theta \frac{1}{2}\|y(t) - y_d(t)\|_H^2 + \frac{\sigma_u}{2}\|u(t)\|_{\mathbb{R}^{N_u}}^2 + \frac{\sigma_w}{2}\|w(t)\|_{\mathbb{R}^{N_w}}^2 \,\mathrm{d}t \qquad \text{(OCP)}$$

on the time interval $\Theta = [0,T]$ subject to the linear parabolic pde constraint

$$\langle \dot{y}(t), \varphi\rangle_{V^\star,V} + \langle \mathcal{A}y(t), \varphi\rangle_{V^\star,V} = \langle \mathcal{B}u(t), \varphi\rangle_{V^\star,V} \qquad \forall \varphi \in V$$
$$\langle y(0), \varphi\rangle_H = \langle y_\circ, \varphi\rangle_H \qquad \forall \varphi \in H$$

and the control and state constraints

$$y_a(t) \le \varepsilon w(t) + (\mathcal{I}y)(t) \le y_b(t) \qquad \& \qquad u_a(t) \le u(t) \le u_b(t),$$

with the operators $\mathcal{B} : L^2(\Theta, \mathbb{R}^{N_u}) \to L^2(\Theta, H)$ and $\mathcal{I} : L^2(\Theta, H) \to L^2(\Theta, \mathbb{R}^{N_w})$,

$$(\mathcal{B}u)(t,x) = \sum_{i=1}^{N_u} u_i(t)\chi_i(x), \qquad (\mathcal{I}y)_i(t) = \int_\Omega \pi_i(x)\, y(t,x)\mathrm{d}x.$$

Universität Konstanz

Transformation on pure box constraints

Introducing a transformed penalty $\omega(t) = \varepsilon w(t) + \mathcal{I}y(t)$, we get the equivalent transformed optimal control problem (TOCP)

$$\min_{y,u,\omega} \tilde{J}(y,u,\omega) = \int_\Theta \frac{1}{2}\|y(t) - \hat{y}_d(t)\|_H^2 + \frac{\sigma_u}{2}\|u(t)\|_{\mathbb{R}^{N_u}}^2 + \frac{\sigma_w}{2\varepsilon^2}\|\omega(t) - \mathcal{I}y(t)\|_{\mathbb{R}^{N_w}}^2 \,\mathrm{d}t$$

subject to the *homogeneous* pde

$$\dot{y}(t) + \mathcal{A}y(t) = \mathcal{B}u(t) \qquad \& \qquad y(0) = 0$$

and the *explicit* penalty and control constraints

$$\hat{y}_a(t) \le \omega(t) \le \hat{y}_b(t) \qquad \& \qquad u_a(t) \le u(t) \le u_b(t)$$

where $\hat{y}_d = y_d - \hat{y}$, $\hat{y}_a = y_a - \mathcal{I}\hat{y}$, $\hat{y}_b = y_b - \mathcal{I}\hat{y}$ and $\hat{y}$ solves

$$\dot{\hat{y}}(t) + \mathcal{A}\hat{y}(t) = 0 \qquad \& \qquad \hat{y}(0) = y_\circ.$$

Universität Konstanz

Well-posedness and optimality conditions

THEOREM. Assume that the closed, convex and bounded set

$$\{(y,u,\omega) \mid \dot{y} + \mathcal{A}y = \mathcal{B}u \;\&\; y(0) = 0 \;\&\; u \in [u_a, u_b] \;\&\; \omega \in [\hat{y}_a, \hat{y}_b]\}$$

is nonempty. Then there exists a unique solution $(\bar{y}, \bar{u}, \bar{\omega})$ to (TOCP).

Further, defining the *active and inactive sets*

$$\mathcal{A}_a^u = \{t \mid \tfrac{1}{\sigma_u}\mathcal{B}^\star\bar{p} < u_a\}, \qquad \mathcal{A}_b^u = \{t \mid \tfrac{1}{\sigma_u}\mathcal{B}^\star\bar{p} > u_a\}, \qquad \mathcal{A}_i^u = \Theta\backslash(\mathcal{A}_a^u \cup \mathcal{A}_b^u)$$
$$\mathcal{A}_a^y = \{t \mid \mathcal{I}\bar{y} < \hat{y}_a\}, \qquad \mathcal{A}_b^y = \{t \mid \mathcal{I}\bar{y} > \hat{y}_b\}, \qquad \mathcal{A}_i^y = \Theta\backslash(\mathcal{A}_a^y \cup \mathcal{A}_b^y),$$

the following first-order optimality system (OS) is fulfilled:

$$\dot{\bar{y}}(t) + \mathcal{A}y(t) - \mathcal{B}\bar{u}(t) = 0, \qquad \bar{y}(0) = 0,$$
$$-\dot{\bar{p}}(t) + \mathcal{A}p(t) + \tfrac{\sigma_w}{\varepsilon^2}(\mathcal{I}^\star(\mathcal{I}y(t) - \omega(t))) + (y(t) - \hat{y}_d(t)) = 0, \qquad \bar{p}(T) = 0,$$
$$\bar{u}(t) - \tfrac{1}{\sigma_u}\chi_i^u(t)\mathcal{B}^\star\bar{p}(t) - (\chi_a^u(t)u_a(t) + \chi_b^u(t)u_b(t)) = 0,$$
$$\bar{\omega}(t) - \chi_i^y(t)\mathcal{I}\bar{y}(t) - (\chi_a^y(t)\hat{y}_a(t) + \chi_b^y(t)\hat{y}_b(t)) = 0.$$

Universität Konstanz

Primal-dual active set strategy (PDASS)

Algorithm (Primal-dual active set strategy)

Require: Initial state-adjoint state pair (y^0, p^0).

1: Set $k = 0$
2: **repeat**
3: Calculate the six active and inactive sets with respect to (y^k, p^k).
4: Solve *linear* primal-dual system (OS) with these fix sets to get (y^{k+1}, p^{k+1})
5: Set $k = k + 1$
6: **until** the current and the previous active and inactive sets coincide.
7: Return control $\bar{u} \in L^2(\Theta, \mathbb{R}^{N_u})$ and penalty $\bar{w} = \frac{1}{\varepsilon}(\omega - \mathcal{I}y) \in L^2(\Theta, \mathbb{R}^{N_w})$.

Universität Konstanz

Proper orthogonal decomposition (POD)

Problem: After elimination of u, ω, the discrete linear system (OS) is still of the dimension $2N_tN_x$.

Idea: For $\ell \ll N_x$, find an *optimal* orthonormal system $\psi = (\psi_1, ..., \psi_\ell) \in V^\ell$ such that the projection error of $\bar{y}$ on the space span(ψ) is minimal:

$$\min_{\phi \in V^\ell \text{ ONB}} \int_\Theta \left\| \bar{y}(t) - \sum_{i=1}^{\ell} \langle \bar{y}(t), \phi_i \rangle_V \phi_i \right\|_V^2 \mathrm{d}t. \qquad \text{(POD)}$$

Realization: Perform an eigenvalue decomposition of a compact, self-adjoint, non-negative operator $\mathcal{R} : V \to V$ which includes the dynamics of the state solution.

Challange: The *optimal* Galerkin ansatz requires the knowledge of the state solution $\bar{y}$ which is not available.

Universität Konstanz

Proper orthogonal decomposition (POD)

Theorem. (Continuous version) Let $y \in C(0, T; V)$ be an *arbitrary* state and let $(\lambda_i, \psi_i)_{i\in\mathbb{N}}$ be an eigenvalue decomposition of

$$\mathcal{R}(y) : V \to V, \qquad \mathcal{R}(y)\varphi = \int_\Theta \langle y(t), \varphi \rangle_V y(t) \, \mathrm{d}t.$$

with $\lambda_i \geq \lambda_{i+1}$ for all $i \in \mathbb{N}$.

Then $(\psi_i)_{i\in\mathbb{N}}$ is a complete orthonormal system in V and the *rank-ℓ POD basis* $\psi^\ell = (\psi_1, ..., \psi_\ell)$ is a solution to (POD).

A priori estimate: The projection error of y on $V^\ell = \text{span}(\psi)$ fulfills

$$\int_\Theta \left\| y(t) - \sum_{i=1}^{\ell} \langle y(t), \phi_i \rangle_V \phi_i \right\|_V^2 = \sum_{i=\ell+1}^{\infty} \lambda_i.$$

Universität Konstanz

Proper orthogonal decomposition (POD)

Discrete POD: Let $(t_1, \dots, t_{N_t}) \subseteq \Theta$ be a time discretization scheme with stepsize Δt and let $\varphi = (\varphi_1, \dots, \varphi_{N_x}) \subseteq V$ be a finite element basis with corresponding weights matrix $\mathrm{X} = (\langle \varphi_i, \varphi_j \rangle_V)$. Let $\mathrm{Y} \in \mathbb{R}^{N_x \times N_t}$ be the coefficient matrix of a state

$$y^{\mathrm{FE}}(t_j, x) = \sum_{i=1}^{N_x} \mathrm{Y}_{ij} \varphi_i(x).$$

Then the coefficient matrix of a rank-ℓ POD basis $\psi \in \mathbb{R}^{N_x \times \ell}$ for y^{FE} is given by the discrete eigenvalue problem

$$\Delta t \mathrm{Y} \mathrm{Y}^{\mathrm{T}} \mathrm{X} \psi_l = \lambda_l \psi_l$$

and the l-th POD element in $V^{N_x} = \mathrm{span}(\varphi)$ is represented by

$$\psi_l = \sum_{i=1}^{N_x} \psi_{il} \varphi_i.$$

With $\mathrm{Y}^\ell = \mathrm{Y}^{\mathrm{T}} \mathrm{X} \psi \in \mathbb{R}^{\ell \times N_t}$, the POD approximation y^{POD} of y^{FE} is given by

$$y^{\mathrm{POD}}(t_j, x) = \sum_{l=1}^{\ell} \mathrm{Y}^\ell_{lj} \psi_l(x).$$

Universität Konstanz

Reduced order model (ROM)

1. **ROM components:**
   - $\mathrm{M} = (\langle \psi_i, \psi_j \rangle_H) \in \mathbb{R}^{\ell \times \ell}$ and $\mathrm{A} = (\langle \mathcal{A} \psi_i, \psi_j \rangle_{V^\star, V}) \in \mathbb{R}^{\ell \times \ell}$.
   - $\mathrm{B} = (\langle \mathcal{B}_j^\star, \psi_i \rangle_{V^\star, V}) \in \mathbb{R}^{\ell \times N_u}$ and $\mathrm{I} = (\mathcal{I}_j \psi_i) \in \mathbb{R}^{N_w \times \ell}$.
   - $\hat{\mathrm{y}}_d(t) = (\langle \hat{y}_d(t), \psi_i \rangle_H) \in \mathbb{R}^\ell$.
2. **ROM system:**

$$\begin{aligned} \mathrm{M}\dot{\mathrm{y}}(t) + \mathrm{A}\mathrm{y}(t) - \mathrm{B}\mathrm{u}(t) &= 0, & \mathrm{y}(0) &= 0, \\ -\mathrm{M}\dot{\mathrm{p}}(t) + \mathrm{A}\mathrm{p}(t) + \tfrac{\sigma_w}{\varepsilon^2}(\mathrm{I}^{\mathrm{T}}(\mathrm{I}\mathrm{y}(t) - \omega(t))) + (\mathrm{M}\mathrm{y}(t) - \hat{\mathrm{y}}_d(t)) &= 0, & \mathrm{p}(T) &= 0, \\ \mathrm{u}(t) - \tfrac{1}{\sigma_u}\chi_i^u(t)\mathrm{B}^{\mathrm{T}}\mathrm{p}(t) - (\chi_i^u(t)u_a(t) + \chi_b^u(t)u_b(t)) &= 0, \\ \omega(t) - \chi_i^y(t)\mathrm{I}\mathrm{y}(t) - (\chi_a^y(t)\hat{y}_a(t) + \chi_b^y(t)\hat{y}_b(t)) &= 0. \end{aligned}$$

3. **ROM expansions:**

$$y^\ell(t) = \sum_{l=1}^{\ell} \mathrm{y}_l(t)\psi_l, \quad p^\ell(t) = \sum_{l=1}^{\ell} \mathrm{p}_l(t)\psi_l, \quad u^\ell = \mathrm{u}, \quad \omega^\ell = \omega$$

Universität Konstanz

Reduced order model (ROM)

A posteriori error bound: Let (u, ω) be any suboptimal control-penalty pair. Then there exists some computable $\zeta \in L^2(\Theta, \mathbb{R}^{N_u} \times \mathbb{R}^{N_w})$ such that

$$\int_\Theta \|u(t) - \bar{u}(t)\|^2_{\mathbb{R}^{N_u}} + \|\omega(t) - \bar{\omega}(t)\|^2_{\mathbb{R}^{N_w}} \, \mathrm{d}t \leq \frac{1}{\min(\sigma_u^2, \sigma_w^2)} \int_\Theta \|\zeta(t)\|^2_{\mathbb{R}^{N_u} \times \mathbb{R}^{N_w}} \, \mathrm{d}t.$$

In our numerical tests, the evaluation of this error estimator is *cheap* (compared to the effort of solving the optimization problem). Problems may arise here in case of *nonlinear* pdes.

In many applications, the error bounds are *sharp*, i.e. the error bound has the same order as the error itself.

Universität Konstanz

Reduced order model (ROM)

Algorithm (Model reduction with iterative POD basis updates (IPOD))

Require: Initial control-penalty pair $(u^{(0)}, \omega^{(0)})$, POD basis rank ℓ, desired exactness ε, maximal iteration number $k_{\max}$.

1: Set $k = 0$
2: **repeat**
3: Solve the full state equation for $y^{(k)}$ and the full adjoint state equation for $p^{(k)}$.
4: Solve the POD eigenvalue problem for the rank-ℓ basis $\psi^{(k)}$.
5: Choose PDASS initialization $(\mathrm{y}, \mathrm{p}) = ((\langle y^{(k)}, \psi_l \rangle_H), (\langle p^{(k)}, \psi_l \rangle_H))$ and provide the PDASS algorithm to solve the ROM system; get feedback $(u^{(k)}, \omega^{(k)})$.
6: Set $k = k + 1$
7: **until** $\mathrm{Aposti}(u^{(k)}, \omega^{(k)}) < \varepsilon$ or $k > k_{\max}$.
8: Return control $u^{(k)} \in L^2(\Theta, \mathbb{R}^{N_u})$ and penalty $w^{(k)} = \frac{1}{\varepsilon}(\omega^{(k)} - \mathcal{I}y) \in L^2(\Theta, \mathbb{R}^{N_w})$.

Universität Konstanz

Optimality system proper orthogonal decomp. (OSPOD)

The optimal state required to determine the POD basis is known *implicitly*:

$$\min_{y,u,\omega,\psi} \tilde{\mathrm{J}}(\mathrm{y},u,\omega,\psi) = \int_\Theta \frac{1}{2}\Big\|\sum_{l=1}^{\ell} \mathrm{y}_l\psi_l - \hat{\mathrm{y}}_d\Big\|_H^2 + \frac{\sigma_u}{2}\|u\|^2_{\mathbb{R}^{N_u}} + \frac{\sigma_w}{2\varepsilon^2}\|\omega - \mathrm{I}(\psi)\mathrm{y}\|^2_{\mathbb{R}^{N_w}}\,\mathrm{d}t$$

subject to the *two* state equations

$$\dot{y} + \mathcal{A}y = \mathcal{B}u, \qquad y(0) = 0, \tag{1}$$

$$\mathrm{M}(\psi)\dot{\mathrm{y}} + \mathrm{A}(\psi)\mathrm{y} = \mathrm{B}(\psi)u, \qquad \mathrm{y}(0) = 0, \tag{2}$$

the POD eigenvalue problem

$$\mathcal{R}(y)\psi_l - \lambda_l\psi_l = 0, \qquad \|\psi_l\|_V^2 = 1 \tag{3}$$

and the penalty and control constraints

$$\hat{y}_a(t) \le \omega(t) \le \hat{y}_b(t) \qquad \& \qquad u_a(t) \le u(t) \le u_b(t).$$

Universität Konstanz

Optimality system proper orthogonal decomp. (OSPOD)

The corresponding dual system is similar to the optimality equations above:

$$-\dot{p} + \mathcal{A}p + \sum_{l=1}^{\ell}\langle y,\psi_l\rangle_V\mu_l + \sum_{l=1}^{\ell}\langle y,\mu_l\rangle_V\psi_l = 0, \tag{4}$$

$$-\mathrm{M}\dot{\mathrm{p}} + \mathrm{A}\mathrm{p} + \frac{\sigma_w}{\varepsilon^2}\mathrm{I}^{\mathrm{T}}(\mathrm{I}\mathrm{y} - \omega) + (\mathrm{M}\mathrm{y} - \hat{\mathrm{y}}_d) = 0, \tag{5}$$

$$u - \frac{1}{\sigma_u}\chi_i^u(\mathcal{B}^\star p + \mathrm{B}^{\mathrm{T}}\mathrm{p}) - (\chi_a^u u_a + \chi_b^u u_b) = 0, \tag{6}$$

$$\omega - \chi_i^y\mathrm{I}\mathrm{y} - (\chi_a^y\hat{y}_a + \chi_b^y\hat{y}_b) = 0, \tag{7}$$

$$\mathcal{R}(y)\mu_l - \lambda_l\mu_l + \mathcal{N}(\mathrm{y},\mathrm{p},u,\omega,\psi) = 0. \tag{8}$$

with some nonlinear term $\mathcal{N}$ arising by the ψ-differential and the active sets

$$\mathcal{A}_a^u = \{t \mid \tfrac{1}{\sigma_u}(\mathcal{B}^\star\bar{p} + \mathrm{B}^{\mathrm{T}}\mathrm{p}) < u_a\}, \qquad \mathcal{A}_b^u = \{t \mid \tfrac{1}{\sigma_u}(\mathcal{B}^\star\bar{p} + \mathrm{B}^{\mathrm{T}}\mathrm{p}) > u_a\},$$

$$\mathcal{A}_a^y = \{t \mid \mathrm{I}\mathrm{y} < \hat{y}_a\}, \qquad \mathcal{A}_b^y = \{t \mid \mathrm{I}\mathrm{y} > \hat{y}_b\}.$$

Universität Konstanz

Optimality system proper orthogonal decomp. (OSPOD)

Algorithm (Optimality System POD (OSPOD))

Require: Initial control-penalty pair (u, ω), POD basis rank ℓ, desired exactness ε.
1: **repeat**
2: Solve the full state equation (1) for y.
3: Solve the eigenvalue problem (3) for ψ.
4: Solve the ROM problem (2), (5), (6), (7) for $(\mathrm{y}, \mathrm{p}, u, \omega)$.
5: Solve the "linearized eigenvalue problem" (8) for μ.
6: Solve the full adjoint equation (4) for p.
7: Provide a descent step in direction $-\sigma_u u + \mathcal{B}^\star p + \mathrm{B}^\mathrm{T}\mathrm{p}$ for u
8: Provide a descent step in direction $-\frac{\sigma_w}{\varepsilon^2}(\omega - \mathrm{Iy})$ for ω.
9: **until** $\mathrm{Aposti}(u, \omega) < \varepsilon$.
10: Return control $u \in L^2(\Theta, \mathbb{R}^{N_u})$ and penalty $w = \frac{1}{\varepsilon}(\omega - \mathrm{Iy}) \in L^2(\Theta, \mathbb{R}^{N_w})$.

Universität Konstanz

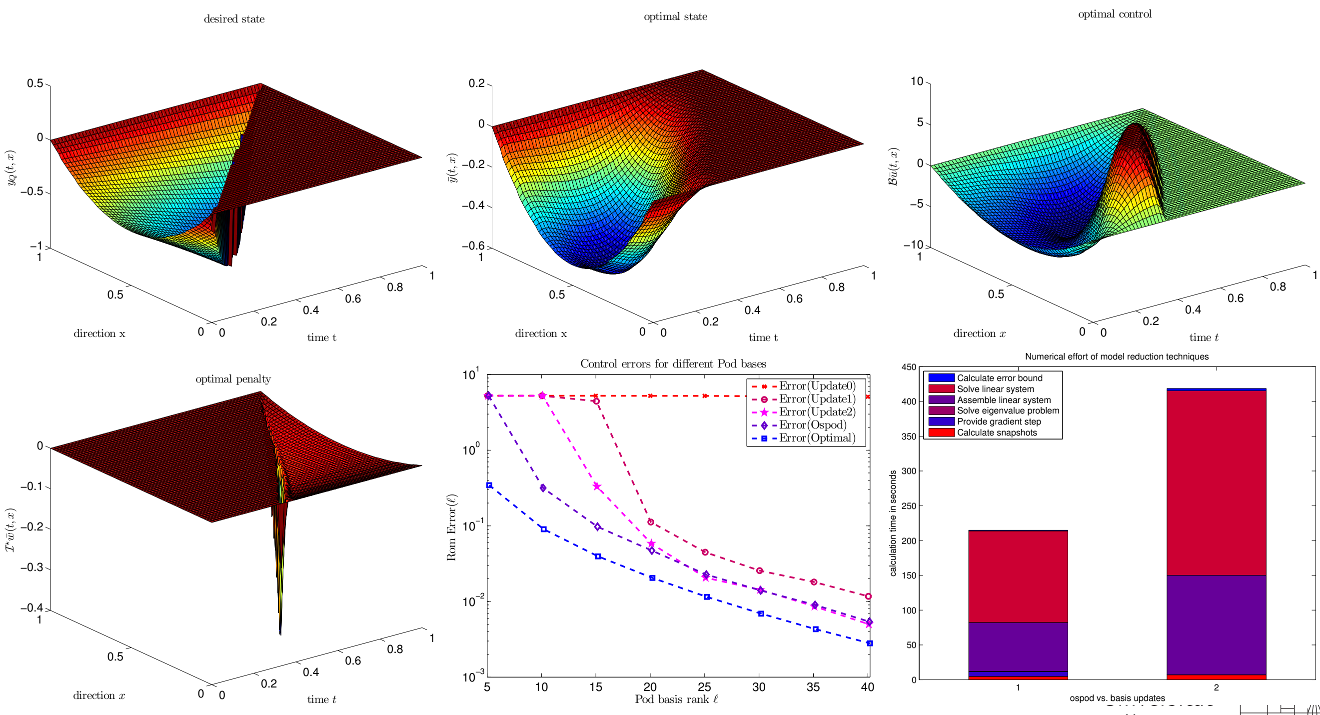
Conclusion: POD & OSPOD

POD: Use a *problem specific* Galerkin ansatz with respect to some reference trajectory y.

OSPOD: Use the trajectory of the optimal state $\bar{y}$ to get the *optimal* Galerkin basis.

Universität Konstanz

Numerical experiments



Konstanz

References

Grimm, E., Gubisch, M. & Volkwein, S.: *Numerical Analysis of OSPOD for Constrained Optimal Control*.
to appear soon.

Gubisch, M. & Volkwein, S.: *POD Reduced-Order Modelling for PDE Constrained Optimization*.
to appear soon, preprint available on KOPS.
URL `http://nbn-resolving.de/urn:nbn:de:bsz:352-250378`.

Gubisch, M. & Volkwein, S.: *POD a-posteriori error analysis for optimal control Problems with mixed control-state constraints*.
Comp. Opt. Appl., **vol. 58**, no. 3: pp. 619–644, 2014.

Hintermüller, M., Kopacka, I. & Volkwein, S.: *Mesh-independence and preconditioning for solving parabolic control problems with mixed control-state constraints*.
ESAIM: COVC, **vol. 15**: pp. 626–652, 2009.

Hinze, M. & Volkwein, S.: *Error estimates for abstract linear-quadratic optimal control problems using proper orthogonal decomposition*.
Comput. Optim. Appl., **vol. 39**: pp. 319–345, 2008.

Kunisch, K. & Volkwein, S.: *Proper orthogonal decomposition for optimality systems*.
ESAIM M2AN, **vol. 42**: pp. 1–23, 2008.

Tröltzsch, F. & Volkwein, S.: *POD a-posteriori error estimates for linear-quadratic optimal control problems*.
Comput. Optim. Appl., **vol. 44**: pp. 319–345, 2008.

Universität
Konstanz