WWU MÜNSTER

Using pyMOR for Reduced Basis PDE-Constrained Optimization

pyMOR School

T. Keil, L. Mechelli, M. Ohlberger, F. Schindler, S. Volkwein

October 9, 2019

Outline

Mathematical background

Model Order Reduction

Live demo

Acknowledgement

- DFG project

Localized Reduced Basis Methods
for PDE-constrained Parameter Optimization

Collaboration

- University of Konstanz: L. **Mechelli**, S. **Volkwein** (Optimization)
- University of Münster: T. **Keil**, F. **Schindler**, M. **Ohlberger** (Model Reduction)

Deutsche Forschungsgemeinschaft
German Research Foundation

Quadratic PDE-Constrained Parameter Optimization

- Quadratic objective functional $\mathcal{J}: V \times \mathcal{P} \to \mathbb{R}$

$$\mathcal{J}(u, \mu) = \Theta(\mu) + j_\mu(u) + k_\mu(u, u)$$

PDE-Constrained Optimization

Find the minimizer

$$\bar{\mu} := \operatorname{argmin}_{\mu \in \mathcal{P}} \hat{\mathcal{J}}(\mu) := \mathcal{J}(u_\mu, \mu), \tag{P.a}$$

subject to $u_\mu \in V$ being the primal state solution of

$$a_\mu(u_\mu, v) = l_\mu(v) \qquad \forall\, v \in V, \tag{P.b}$$

Assumptions:

- real-valued Hilbert space V (e.g. $V = H_0^1(\Omega)$).
- compact Banach space $\mathcal{P}$ (e.g. $\mathcal{P} = \mathbb{R}^p$ for $p \in \mathbb{N}$)
- parameter functional $\Theta \in \mathcal{P}'$.
- $a_\mu, k_\mu: V \times V \to \mathbb{R}$ continuous (and coercive) bilinear forms.
- $l_\mu, j_\mu: V \to \mathbb{R}$ continuous linear functionals
- $a_\mu, l_\mu, j_\mu, k_\mu$ parameter separable.

$$a_\mu(u, v) = \sum_{q=1}^{Q} \theta^q(\mu) a^q(u, v)$$

Quadratic PDE-Constrained Parameter Optimization

- Quadratic objective functional $\mathcal{J}: V \times \mathcal{P} \to \mathbb{R}$

$$\mathcal{J}(u,\mu) = \Theta(\mu) + j_\mu(u) + k_\mu(u,u)$$

PDE-Constrained Optimization

Find the minimizer

$$\bar{\mu} := \operatorname{argmin}_{\mu\in\mathcal{P}} \hat{\mathcal{J}}(\mu) := \mathcal{J}(u_\mu,\mu), \qquad \text{(P.a)}$$

subject to $u_\mu \in V$ being the primal state solution of

$$a_\mu(u_\mu, v) = l_\mu(v) \qquad \forall\, v \in V, \qquad \text{(P.b)}$$

Challenges for optimization:

- efficient evaluation of $\hat{\mathcal{J}}(\mu)$, gradient $\nabla\hat{\mathcal{J}}(\mu)$ or even Hessian $\mathcal{H}\hat{\mathcal{J}}(\mu)$.
  $\Rightarrow$ Efficient evaluation of the primal state equation
- high dimensional parameter space $\mathcal{P}$.
- multiscale/large scale context (localization).

Efficient optimization method using (localized) MOR techniques.

Trust region methods

- Some optimization methods: Gradient descent, Projected Newton Method, BFGS.

Trust region

Approximate

$$\bar{\mu} = \text{argmin}_{\mu \in \mathcal{P}} \hat{\mathcal{J}}(\mu)$$

by solving optimization subproblems

$$\mu_{k+1} := \mu_k + s_k, \qquad \min_{s \in \mathcal{P}} m_k(s), \qquad s.t. \|s\|_{\mathcal{P}} \leq \delta_k,$$

starting with some m_0, δ_0.

Requirements:

- a **model function** $m_k \approx \hat{\mathcal{J}}$, smooth and $\|\nabla m_k\| < \infty$.
- bound on $|m_k(\mu) - \hat{\mathcal{J}}(\mu)|$ for all $\mu \in \mathcal{P}$.

How to choose m_k?

first idea:

$$m_k(s) := \hat{\mathcal{J}}(\mu_k + s) \quad . \tag{3}$$

Trust region methods

- Some optimization methods: Gradient descent, Projected Newton Method, BFGS.

Trust region

Approximate

$$\bar{\mu} = \operatorname{argmin}_{\mu \in \mathcal{P}} \hat{\mathcal{J}}(\mu)$$

by solving optimization subproblems

$$\mu_{k+1} := \mu_k + s_k, \qquad \min_{s \in \mathcal{P}} m_k(s), \qquad s.t. \|s\|_{\mathcal{P}} \leq \delta_k,$$

starting with some m_0, δ_0.

Requirements:

- a **model function** $m_k \approx \hat{\mathcal{J}}$, smooth and $\|\nabla m_k\| < \infty$.
- bound on $|m_k(\mu) - \hat{\mathcal{J}}(\mu)|$ for all $\mu \in \mathcal{P}$.

How to choose m_k?

first idea:

$$m_k(s) := \hat{\mathcal{J}}(\mu_k + s) + \nabla \hat{\mathcal{J}}(\mu_k + s)\mu_k + \dots . \qquad (3)$$

First-optimize-then-discretize

primal state	$a_\mu(u_\mu, v) = l_\mu(v)$
primal residual	$r_\mu^{\mathrm{pr}}(u_\mu)[v] = l_\mu(v) - a_\mu(u_\mu, v)$
primal sensitivity	$a_\mu(d_{\mu_i} u_\mu, v) = \partial_\mu r_\mu^{\mathrm{pr}}(u_\mu)[v] \cdot e_i$
dual state	$a_\mu(q, p_\mu) = \partial_u \mathcal{J}(u_\mu, \mu)[q] = j_\mu(q) + 2k_\mu(q, u_\mu)$
dual residual	$r_\mu^{\mathrm{du}}(u_\mu, p_\mu)[q] = j_\mu(q) + 2k_\mu(q, u_\mu) - a_\mu(q, p_\mu)$
dual sensitivity	$a_\mu(q, d_{\mu_i} p_\mu) = \partial_\mu r_\mu^{\mathrm{du}}(u_\mu, p_\mu)[q] \cdot e_i + 2k_\mu(q, d_{\mu_i} u_\mu)$
reduced gradient	$\nabla_\mu \hat{\mathcal{J}}(\mu) = \nabla_\mu \mathcal{L}(u_\mu, \mu, p_\mu)$ ($\mathcal{L}$ from Lagrangian ansatz (all-at-once))
reduced Hessian	$\left(\mathcal{H}_\mu \hat{\mathcal{J}}(\mu)\right)_{i,l} = d_{\mu_l}\left(\nabla_\mu \mathcal{L}(u_\mu, \mu, p_\mu)\right)_i$

Aim: Use MOR for the model function, e.g.

$$m_k(s) := \hat{\mathcal{J}}_N^k(\mu_k + s).$$

Model Order Reduction with Reduced Basis Methods

Discretization

- computational grid τ_h, approximate space $V_h \approx V$ (e.g. CG FE)
- approximations $j_{h,\mu}$, $k_{h,\mu}$, $l_{h,\mu}$, $a_{h,\mu}$, $\mathcal{J}_h$ and $\hat{\mathcal{J}}_h$
- negligible discretization error (e.g. $\|u_\mu - u_{h,\mu}\|$).

Model Order Reduction

- problem adapted primal and dual RB spaces $V_N^{\mathrm{pr}}, V_N^{\mathrm{du}} \subset V_h$.
- problem adapted sensitivity RB spaces $V_N^{\mathrm{pr},d_{\mu_i}}, V_N^{\mathrm{du},d_{\mu_i}} \subset V_h$, for all i.
- Galerkin projections of $j_{h,\mu}$, $k_{h,\mu}$, $l_{h,\mu}$, $a_{h,\mu}$ onto $V_N^{\mathrm{pr/du}}$

- RB reduced solutions $u_{N,\mu}$, $p_{N,\mu}$ and $d_{\mu_i} u_{N,\mu}$ and $d_{\mu_i} p_{N,\mu}$.
- RB reduced cost functional $\hat{\mathcal{J}}_N(\mu) = \mathcal{J}(u_{N,\mu}, \mu)$ and Hessian $\mathcal{H}\hat{\mathcal{J}}_N(\mu)$.

Challenge: Offline-Online decomposition for e.g. $\nabla \hat{\mathcal{J}}_N(\mu)$ and $\mathcal{H}\hat{\mathcal{J}}_N(\mu)$)

Reduced quantities

Note: $p_{N,\mu}$ and $d_{\mu_i}u_{N,\mu}$ and $d_{\mu_i}p_{N,\mu}$ are no Galerkin projections.

- reduced primal state

$$a_\mu(u_{N,\mu}, v) = l_\mu(v)$$

- reduced dual state

$$a_\mu(q, p_{N,\mu}) = \partial_u \mathcal{J}(u_{N,\mu}, \mu)[q]$$

- reduced primal sensitivities

$$a_\mu(d_{\mu_i}u_{N,\mu}, v) = \partial_\mu r_\mu^{\mathrm{pr}}(u_{N,\mu})[v] \cdot e_i$$

- reduced dual sensitivities

$$a_\mu(q, d_{\mu_i}p_{N,\mu}) = \partial_\mu r_\mu^{\mathrm{du}}(u_{N,\mu}, p_{N,\mu})[q] \cdot e_i + 2k_\mu(q, d_{\mu_i}u_{N,\mu})$$

A posteriori error estimation

$$
\begin{aligned}
\|u_{h,\mu} - u_{N,\mu}\|_h &\leq \Delta_{\mathrm{pr}}(\mu) && \approx \|r_{h,\mu}^{\mathrm{pr}}(u_{N,\mu})\|_h \\
\|p_{h,\mu} - p_{N,\mu}\|_h &\leq \Delta_{\mathrm{du}}(\mu) && \approx \Delta_{\mathrm{pr}}(\mu) + \|r_{h,\mu}^{\mathrm{du}}(u_{N,\mu}, p_{N,\mu})\|_h \\
|\hat{\mathcal{J}}_h(\mu) - \hat{\mathcal{J}}_N(\mu)| &\leq \Delta_{\hat{\mathcal{J}}}(\mu) && \approx \Delta_{\mathrm{pr}}(\mu)(\|u_{N,\mu}\|_h + \Delta_{\mathrm{pr}}(\mu)) \\
\left|(\nabla_\mu \hat{\mathcal{J}}_h(\mu) - \nabla_\mu \hat{\mathcal{J}}_N(\mu))_i\right| &\leq \Delta_{\nabla_\mu \hat{\mathcal{J}}_i}(\mu) && \approx \Delta_{\mathrm{pr}}(\mu)\|u_{N,\mu}\|_h + \Delta_{\mathrm{pr}}(\mu)\|p_{N,\mu}\|_h + \Delta_{\mathrm{du}}(\mu)\|u_{N,\mu}\|_h \\
& && \quad + \Delta_{\mathrm{pr}}(\mu)\,\Delta_{\mathrm{du}}(\mu) + \Delta_{\mathrm{pr}}^2(\mu) \\
\|d_{\mu_i} u_{h,\mu} - d_{\mu_i} u_{N,\mu}\|_h &\leq \Delta_{d_{\mu_i}pr}(\mu) && \approx \Delta_{\mathrm{pr}}(\mu) + \|r_{h,\mu}^{\mathrm{pr},d_{\mu_i}}(u_{N,\mu}, d_{\mu_i} u_{N,\mu})\|_h \\
\|d_{\mu_i} p_{h,\mu} - d_{\mu_i} p_{N,\mu}\| &\leq \Delta_{d_{\mu_i}\mathrm{du}}(\mu) && \approx \Delta_{\mathrm{pr}}(\mu) + \Delta_{\mathrm{du}}(\mu) + \Delta_{d_{\mu_i}\mathrm{pr}}(\mu) + \|r_{h,\mu}^{\mathrm{du},d_{\mu_i}}(u_{N,\mu}, p_{N,\mu}, d_{\mu_i} u_{N,\mu}, d_{\mu_i} p_{N,\mu})\|_h \\
\left|(\mathcal{H}_\mu \hat{\mathcal{J}}_h(\mu) - \tilde{\mathcal{H}} \hat{\mathcal{J}}_N(\mu))_{i,l}\right| &\leq \Delta_{\mathcal{H}}(\mu) && \approx \Delta_{\mathrm{pr}}(\mu) + \Delta_{\mathrm{pr}}^2(\mu) + \Delta_{d_{\mu_i}\mathrm{pr}}(\mu) + \Delta_{d_{\mu_i}\mathrm{pr}}(\mu)\Delta_{\mathrm{pr}}(\mu) + \Delta_{\mathrm{du}}(\mu) \\
& && \quad + \Delta_{d_{\mu_i}\mathrm{du}}(\mu) + \Delta_{\mathrm{pr}}(\mu)\Delta_{\mathrm{du}}(\mu) + \Delta_{d_{\mu_i}\mathrm{du}}(\mu)\Delta_{\mathrm{pr}}(\mu) + \Delta_{d_{\mu_i}\mathrm{pr}}(\mu)\Delta_{\mathrm{du}}(\mu)
\end{aligned}
$$

Further details and future work

How to construct RB spaces?

- Add every parameter that is picked in the optimization algorithm.
- construct a space around a starting parameter (”global greedy” vs. ”greedy in a box”)
- Use only two unique bases for primal and dual RB spaces.
- Use all estimators for the greedy algorithm.

Localization:

- Replace discretization by localized discretization (e.g. DG FEM).
- No global FEM solves required.
- Localized error estimations.
- Adaptive online enrichment (p-like refinement).

live demo