Structured $\mathcal{L}_2$-Optimal Parametric Model Order Reduction

P. Mlinarić[1] and S. Gugercin[1,2]

[1]*Department of Mathematics, Virginia Tech, Blacksburg, VA 24061, USA*
[2]*Division of Computational Modeling and Data Analytics, Academy of Data Sciences, Virginia Tech, Blacksburg, VA 24061, USA*

Consider the full-order model (FOM) and reduced-order model (ROM)

$$\text{FOM: } \begin{cases} \mathcal{A}(\mathsf{p})x(\mathsf{p}) = \mathcal{B}(\mathsf{p}), \\ \quad y(\mathsf{p}) = \mathcal{C}(\mathsf{p})x(\mathsf{p}), \end{cases} \quad (1) \qquad \text{and} \qquad \text{ROM: } \begin{cases} \widehat{\mathcal{A}}(\mathsf{p})\widehat{x}(\mathsf{p}) = \widehat{\mathcal{B}}(\mathsf{p}), \\ \quad \widehat{y}(\mathsf{p}) = \widehat{\mathcal{C}}(\mathsf{p})\widehat{x}(\mathsf{p}), \end{cases} \quad (2)$$

where $\mathsf{p} \in \mathcal{P} \subseteq \mathbb{C}^d$ is the parameter, $x(\mathsf{p}) \in \mathbb{C}^{n \times n_i}$ is the state, $y(\mathsf{p}) \in \mathbb{C}^{n_o \times n_i}$ is the output, $\widehat{x}(\mathsf{p}) \in \mathbb{C}^{r \times n_i}$ is the reduced state for some $r \ll n$, and $\widehat{y}(\mathsf{p}) \in \mathbb{C}^{n_o \times n_i}$ is the approximate output. Therefore, $\mathcal{A}(\mathsf{p}) \in \mathbb{C}^{n \times n}$, $\mathcal{B}(\mathsf{p}) \in \mathbb{C}^{n \times n_i}$, $\mathcal{C}(\mathsf{p}) \in \mathbb{C}^{n_o \times n}$, $\widehat{\mathcal{A}}(\mathsf{p}) \in \mathbb{C}^{r \times r}$, $\widehat{\mathcal{B}}(\mathsf{p}) \in \mathbb{C}^{r \times n_i}$, and $\widehat{\mathcal{C}}(\mathsf{p}) \in \mathbb{C}^{n_o \times r}$. Our goal is to find a ROM (2) that is an $\mathcal{L}_2$-optimal approximation of the form

$$\widehat{\mathcal{A}}(\mathsf{p}) = \sum_{i=1}^{q_{\widehat{\mathcal{A}}}} \widehat{\alpha}_i(\mathsf{p})\widehat{A}_i, \quad \widehat{\mathcal{B}}(\mathsf{p}) = \sum_{j=1}^{q_{\widehat{\mathcal{B}}}} \widehat{\beta}_j(\mathsf{p})\widehat{B}_j, \quad \widehat{\mathcal{C}}(\mathsf{p}) = \sum_{k=1}^{q_{\widehat{\mathcal{C}}}} \widehat{\gamma}_k(\mathsf{p})\widehat{C}_k, \tag{3}$$

where $q_{\widehat{\mathcal{A}}}, q_{\widehat{\mathcal{B}}}, q_{\widehat{\mathcal{C}}}$ are positive integers; $\widehat{\alpha}_i, \widehat{\beta}_j, \widehat{\gamma}_k \colon \mathcal{P} \to \mathbb{C}$ are given measurable functions; and $\widehat{A}_i \in \mathbb{R}^{r \times r}$, $\widehat{B}_j \in \mathbb{R}^{r \times n_i}$, and $\widehat{C}_k \in \mathbb{R}^{n_o \times r}$. By "$\mathcal{L}_2$-optimality", we mean that we seek a ROM that minimizes

$$\mathcal{J}(\widehat{y}) = \|y - \widehat{y}\|_{\mathcal{L}_2(\mathcal{P},\mu)}^2 = \int_{\mathcal{P}} \|y(\mathsf{p}) - \widehat{y}(\mathsf{p})\|_{\mathrm{F}}^2 \,\mathrm{d}\mu(\mathsf{p}), \tag{4}$$

where μ is a measure over the parameter set $\mathcal{P}$.
We first derive the gradients $\nabla_{\widehat{A}_i}\mathcal{J}$, $\nabla_{\widehat{B}_j}\mathcal{J}$, and $\nabla_{\widehat{C}_k}\mathcal{J}$ of the squared $\mathcal{L}_2$ error (4) (with respect to the reduced-order matrices), which then leads to a gradient-based optimization method for structured model order reduction (MOR) of parametric problems. We illustrate that these gradients can be computed using only output values, and thus the optimization algorithm can be performed purely in a data-driven manner based on the samples of the output without access to full-order operators.
We show that by appropriately defining the measure μ and the parameter set $\mathcal{P}$, our formulation in (1)–(4) covers a wide range of problems such as parametric stationary problems, e.g., arising from discretization of parametric partial differential equations, as well as (parametric) linear-time invariant systems and discretized least-squares fitting.
We also show that this unifying framework recovers the well-known optimality conditions for $\mathcal{H}_2$ and $\mathcal{H}_2 \otimes \mathcal{L}_2$-optimal MOR for dynamical systems. Furthermore, we develop interpolatory necessary conditions for $\mathcal{L}_2$-optimal MOR of a class of parametric stationary problems; more precisely, we show that if $\widehat{y}(\mathsf{p}) = \sum_{i=1}^{r} \frac{c_i b_i^{\mathrm{T}}}{\mathsf{p} - \lambda_i}$ is an $\mathcal{L}_2$-optimal ROM (with some additional assumptions), then we have

$$Y(\lambda_i)b_i = \widehat{Y}(\lambda_i)b_i, \quad c_i^{\mathrm{T}} Y(\lambda_i) = c_i^{\mathrm{T}} \widehat{Y}(\lambda_i), \quad c_i^{\mathrm{T}} Y'(\lambda_i)b_i = c_i^{\mathrm{T}} \widehat{Y}'(\lambda_i)b_i,$$

for $i = 1, 2, \ldots, r$, where Y and $\widehat{Y}$ are transformed outputs related to y and $\widehat{y}$. Finally, we discuss MOR methods based on (Petrov-)Galerkin projection and whether $\mathcal{L}_2$-optimal ROMs are necessarily of such type.
We illustrate the theory via various numerical examples and compare our framework to standard projection-based approaches.