Beyond Petrov-Galerkin projection by using « multi-space » priors

Cédric Herzet
Inria, France

Joint work with M. Diallo & P. Héas
The target problem...

Find \( h^* \in \mathcal{H} \) such that \( a(h^*, h) = b(h) \quad \forall h \in \mathcal{H} \)

where \( \mathcal{H} \) is a Hilbert space (\( \langle \cdot, \cdot \rangle \) and \( \| \cdot \| \))
\( a : \mathcal{H} \times \mathcal{H} \rightarrow \mathbb{R} \) is a bilinear operator
\( b : \mathcal{H} \rightarrow \mathbb{R} \) is a linear operator
... and its Petrov-Galerkin approximation

Find $\hat{h}_{PG} \in V_n$ such that $a(\hat{h}_{PG}, h) = b(h) \quad \forall h \in Z_n$

where $V_n \subset \mathcal{H}$, $Z_n \subset \mathcal{H}$ are $n$-dimensional subspaces
The precision of Petrov-Galerkin can be quantified by an « instance optimal property »

$$\| h^* - \hat{h}_{PG} \| \leq C(V_n, Z_n) \text{dist}(h^*, V_n),$$
Standard methods constructing $V_n$ often return a set of subspaces and their «widths»

Standard outputs of methods constructing $V_n$:

$$V_0 \subset V_1 \subset \ldots \subset V_n, \quad \dim(V_k) = k$$

such that

$$\text{dist}(h^*, V_k) \leq \hat{\epsilon}_k, \quad k = 0 \ldots n.$$ 

E.g., «reduced basis» methods
The Petrov-Galerkin projection discards most of the available information.

Standard outputs of methods constructing $V_n$:

$$\mathbf{\times} \subset \mathbf{\times} \subset \ldots \subset V_n, \quad \dim(V_k) = k$$

such that

$$\operatorname{dist}(h^*, \mathbf{\times}) \leq \hat{\varepsilon}_k, \quad k = 0 \ldots n.$$ 

*E.g.*, « reduced basis » methods
Can we use this information to improve the projection process?
The Petrov-Galerkin projection can be reformulated as a variational problem

\[
\hat{h}_{PG} = \arg \min_{h \in V_n} \sum_{j=1}^{n} (b_j - \langle a_j, h \rangle)^2
\]

where

\[
\text{span} \left( \{ z_j \}_{j=1}^{n} \right) = Z_n
\]

\( a_j \) is the Riesz’s representer of \( a(\cdot, z_j) \)

\( b_j = b(z_j) \)
The proposed « multi-space » decoder adds new constraints to the variational problem

\[ \hat{h}_{MS} = \arg \min_{h \in V_n} \sum_{j=1}^{n} \left( b_j - \langle a_j, h \rangle \right)^2, \]

subject to \( \text{dist}(h, V_k) \leq \hat{\varepsilon}_k, \quad k = 0 \ldots n - 1. \)

See [Binev et al., SIAM JUQ 17] for a related work.
A graphical representation of the problem

\[ n = 2 \]

\[ V_k = \text{span} \left( \{ \mathbf{v}_i \}_{i=1}^k \right) \]
A graphical representation of the problem

\[ V_k = \text{span} \left( \{v_i\}_{i=1}^k \right) \]

The shape of the iso-contours depends on \( \mathbf{G} = [\langle \mathbf{a}_i, \mathbf{v}_j \rangle]_{i,j} \)
A graphical representation of the problem

\[ h \cdot P_{V_n} = h_{MS} \]

\[ \text{span}(V_n) \]

\{ h : \text{dist}(h, V_0) \leq \hat{\varepsilon}_0 \}

\[ n = 2 \]

\[ V_k = \text{span} \left( \{ v_i \}_{i=1}^k \right) \]

The shape of the iso-contours depends on \( G = [\langle a_i, v_j \rangle]_{i,j} \)
A graphical representation of the problem

The shape of the iso-contours depends on $G = [\langle a_i, v_j \rangle]_{i,j}$

The feasibility region depends on $\{V_k\}_{k=1}^n$ and $\{\hat{e}_k\}_{k=1}^n$
Can we give some guarantee on the performance of the « multi-space » decoder?
Instance optimality properties

- **Petrov-Galerkin**: \( \| h^* - \hat{h}_{PG} \| \leq C(G) \text{dist}(h^*, V_n) \),

- **« Multi-space » decoder**:

\[
\| h^* - \hat{h}_{MS} \| \leq \left( \sum_{j=\ell+1}^{n} \delta_j^2 + \rho \delta_{\ell}^2 + (\text{dist}(h^*, V_n))^2 \right)^{\frac{1}{2}}
\]

where \( \ell \) and \( \delta_j \)'s are “easily-computable” quantities only depending on \( G = [\langle a_i, v_j \rangle]_{i,j}, \{\hat{\epsilon}_k\}_{k=1}^{n-1} \) and \( \{\text{dist}(h^*, V_k)\}_{k=1}^{n} \).
Particularization of the instance optimality bound to some examples
Particularization of the instance optimality bound to some examples

\[ \hat{e}_j = \begin{cases} 
1 & j = 0 \ldots n - 3, \\
\epsilon^{1/2} & j = n - 2, n - 1, \\
\epsilon & j = n,
\end{cases} \]

\[ \text{dist}(h^*, V_k) = \hat{e}_j \]

\[ \sigma_j = \text{sing. values of } G \]

\[ = \begin{cases} 
1 & j = 1 \ldots n - 3, \\
\epsilon^{1/2} & j = n - 2, n - 1, \\
\epsilon & j = n.
\end{cases} \]

\[ \| h^* - \hat{h}_{PG} \| \leq 1 \]
\[ \| h^* - \hat{h}_{MS} \| \leq 3\epsilon^{1/2} \]
Quid of the computational complexity?
PG projection can be carried out efficiently with a complexity $O(n^2)$ per iteration

$$\hat{h}_{\text{PG}} = \arg \min_{h \in V_n} \sum_{j=1}^{n} (b_j - \langle a_j, h \rangle)^2$$

« Least square » problem: can be solved efficiently via gradient-based methods with a complexity $O(n^2)$ per iteration.
Our decoder can also be implemented with a complexity $\mathcal{O}(n^2)$ per iteration.

Our problem can be rewritten as:

$$
\hat{h}_{MS} = \arg \min_{h \in V_n} \sum_{j=1}^{n} (b_j - \langle a_j, h \rangle)^2,
$$

subject to $\| P_{V_k} (h) \| \leq \hat{\epsilon}_k$, $k = 0 \ldots n - 1$.

We use the « Alternating Directions Method of Multipliers » to solve this convex problem with a complexity $\mathcal{O}(n^2)$ per iteration.
Some results
We consider a parametric mass transfert problem

\[ \mu_1 \Delta h + b(\mu_2) \cdot \nabla h = s \quad \text{in } \Omega \]

\[ \mu_1 \nabla h \cdot n = 0 \quad \text{in } \partial \Omega \]

where

\[ b(\mu_2) = [\cos(\mu_2) \sin(\mu_2)]^T \]

\[ s = \exp \left( - \frac{\|x - m\|_2^2}{2\sigma^2} \right) \]

\[ \mu_1 = [0.03, 0.05] \]

\[ \mu_2 = [0, 2\pi] \]
The multi-space decoder improves the worst-case approximation error over PG
Our contributions

- We exploit additional information in the Petrov-Galerkin projection
- We derive an instance optimal guarantee for our « multi-space » decoder
- We propose an efficient implementation for the « multi-space » decoder

Thank you for your attention.