Goal-oriented error estimation for reduced basis method

Alexandre Janon (Univ. Paris Sud), with:
Maëlle Nodet, Clémentine Prieur (U. Grenoble 1, INRIA)

Nice, July 2013
Context

- $\mu \in \mathcal{P} \subset \mathbb{R}^p$: input parameter.
- We want to compute a model output $s(\mu)$ for many values of $\mu$.
- We suppose that $s$ is a linear functional:
  \[ s(\mu) = l^t u(\mu), \]
  where $u(\mu)$ is the solution of the linear system:
  \[ A(\mu) u(\mu) = f(\mu), \]
  where $A(\mu)$ and $f(\mu)$ are known matrix/vector.
- Typically, the linear system is obtained by discretizing a (linear) PDE given by the physics, and the $u(\mu) \mapsto s(\mu)$ operation is evaluation or mean.
- **Problem**: $u(\mu)$ is of dimension $N \gg 1$.
- In a many-query context, solving the system for every parameter of interest may be too long.
Context (2) – Reduced basis method

- The idea is to project the large system onto a smaller subspace. Given a (well-chosen) matrix $Z$ with $n$ cols and $N$ lines, we look for $\tilde{u}(\mu) \in \mathbb{R}^n$ so that:

$$ (Z^t A(\mu) Z) \tilde{u}(\mu) = Z^t f(\mu). $$

- The system is of dimension $n$. Fine if $n \ll N$.
- If $u(\mu)$ is in the range of $Z$, then the system above is equivalent to the original one:

$$ A(\mu) u(\mu) = f(\mu), $$

and we have $u(\mu) = Z\tilde{u}(\mu)$.
- In many interesting cases, we have methods to choose $Z$ so that

$$ n \ll N \quad \text{and} \quad u(\mu) \approx Z\tilde{u}(\mu) \quad \text{for many} \quad \mu. $$

and so:

$$ \tilde{s}(\mu) = l^t Z\tilde{u}(\mu) \approx l^t u(\mu) = s(\mu). $$

- $\tilde{s}(\mu)$: metamodel.
- Can we quantify the error in this approximation?
Under some hypotheses on the $A(\mu)$ matrix and a norm $\|\cdot\|$ (say, Euclidean norm), the reduced basis comes with an error bound $\epsilon_u(\mu)$:

$$\forall \mu \in \mathcal{P}, \|u(\mu) - Z\tilde{u}(\mu)\| \leq \epsilon_u(\mu)$$

which can be numerically computed efficiently (i.e., with the order of complexity of the computation of $\tilde{u}(\mu)$).

**Question:** Given this bound, can we have an error bound $\epsilon(\mu)$ on $s$:

$$\forall \mu \in \mathcal{P}, |s(\mu) - \tilde{s}(\mu)| \leq \epsilon(\mu)$$

which can be explicitly and efficiently computed?

**Yes,** as the “Lipschitz bound” holds:

$$\forall \mu \in \mathcal{P}, |s(\mu) - \tilde{s}(\mu)| \leq L\epsilon_u(\mu),$$

for:

$$L = \sup_{\|v\|=1} l^t v.$$
Question: can we find a more precise error bound?

The Lipschitz bound is optimal amongst the bounds which depend on (a bound on) \( \| u(\mu) - A(\mu)\tilde{u}(\mu) \| \).

Our improved bound has to depend on something else...

Contents of the talk:
- Description of the proposed bound
- Further improvement: correction of the output
- Numerical examples and comparisons

Starting point

- Remember: $A(\mu) Z \tilde{u}(\mu) \approx f(\mu)$.
- The bound $\varepsilon^u(\mu)$ on $\|u(\mu) - \tilde{u}(\mu)\|$ is based on the residual:
  \[ r(\mu) = A(\mu) Z \tilde{u}(\mu) - f(\mu), \]
  and that its norm is efficiently computable.
- We also want to exploit that the (say, Euclidean) scalar products of the residual:
  \[ \langle r(\mu), \phi \rangle \]
  by any vector $\phi$ are also efficiently computable.
- Let $\{\phi_i\}_{i=1,\ldots,N}$ be an orthonormal basis of $\mathbb{R}^N$ (to be choosed later). We have:
  \[ \tilde{s}(\mu) - s(\mu) = \sum_{i \geq 1} \langle w(\mu), \phi_i \rangle \langle r(\mu), \phi_i \rangle, \]
  where $w(\mu)$ is the solution of the \textbf{adjoint} (or \textbf{dual}) problem:
  \[ w(\mu) = A(\mu)^{-t} I, \]
  we set $\phi_i = 0$ for $i > N$. 

Error bound – Two-part decomposition

Let $N \in \mathbb{N}^*$. We have:

$$|\tilde{s}(\mu) - s(\mu)| = \left| \sum_i \langle w(\mu), \phi_i \rangle \langle r(\mu), \phi_i \rangle \right|$$

$$\leq \left| \sum_{i=1}^N \langle w(\mu), \phi_i \rangle \langle r(\mu), \phi_i \rangle \right| + \left| \sum_{i>N} \langle w(\mu), \phi_i \rangle \langle r(\mu), \phi_i \rangle \right|$$

- The first term is to be bounded by a $\mu$-dependent quantity which can be computed efficiently.
- The second term will be:
  - bounded, in probability (with respect to $\mu$), by a $\mu$-independent quantity;
  - (heuristically) minimized by the choice of $\{\phi_i\}_i$. 

Bound – Addressment of the first term

Let:

$$\tau_1(\mu) := \left| \sum_{i=1}^{N} \langle w(\mu), \phi_i \rangle \langle r(\mu), \phi_i \rangle \right|$$

computed to bound

We compute (once for all the values of $\mu$):

$$\beta_{i}^{\text{min}} = \min_{\mu \in \mathcal{P}} D_i(\mu), \quad \beta_{i}^{\text{max}} = \max_{\mu \in \mathcal{P}} D_i(\mu),$$

where:

$$D_i(\mu) = \langle w(\mu), \phi_i \rangle.$$  

(2$N$ optimization problems to solve on $\mathcal{P}$.)

We set:

$$\beta_{i}^{\text{up}}(\mu) = \begin{cases} \beta_{i}^{\text{max}} & \text{if } \langle r(\mu), \phi_i \rangle > 0 \\ \beta_{i}^{\text{min}} & \text{else,} \end{cases} \quad \beta_{i}^{\text{low}}(\mu) = \begin{cases} \beta_{i}^{\text{min}} & \text{if } \langle r(\mu), \phi_i \rangle > 0 \\ \beta_{i}^{\text{max}} & \text{else.} \end{cases}$$

and we have:

$$|\tau_1(\mu)| \leq \max \left( \left| \sum_{i=1}^{N} \langle r(\mu), \phi_i \rangle \beta_i^{\text{low}}(\mu) \right|, \left| \sum_{i=1}^{N} \langle r(\mu), \phi_i \rangle \beta_i^{\text{up}}(\mu) \right| \right) =: T_1(\mu).$$
Bound – Addressment of the second term

Let:

$$\tau_2(\mu) = \left| \sum_{i > N} \langle w(\mu), \phi_i \rangle \langle r(\mu), \phi_i \rangle \right| .$$

Not efficiently computable.

We assume that $\mu$ is a random variable on $\mathcal{P}$, with known distribution.

We want to control $E_\mu [\tau_2(\mu)]$.

We have:

$$E_\mu [\tau_2(\mu)] \leq \frac{1}{2} E_\mu \left( \sum_{i > N} \langle w(\mu), \phi_i \rangle^2 + \sum_{i > N} \langle r(\mu), \phi_i \rangle^2 \right) = \sum_{i > N} \langle G\phi_i, \phi_i \rangle$$

where $G$ is the positive, self-adjoint operator given by:

$$\forall \phi \in X, \quad G\phi = \frac{1}{2} E_\mu (\langle r(\mu), \phi \rangle r(\mu) + \langle w(\mu), \phi \rangle w(\mu)) .$$
Bound – Addressment of the second term (2)

- Recall that:
  \[ E_\mu [\tau_2(\mu)] \leq \sum_{i > N} \langle G\phi_i, \phi_i \rangle. \]

- Let \( \lambda_1 \geq \lambda_2 \geq \ldots \lambda_N \geq 0 \) be the eigenvalues of \( G \), and \( \phi_i^G \) a unitary eigenvector of \( G \) with respect to \( \lambda_i \).
- The RHS is minimized for \( \phi_i = \phi_i^G \) \( \forall i > N \).
- This suggests to choose
  \[ \phi_i = \phi_i^G \forall i \leq N, \]
  so have to the a priori bound on \( \tau_2 \):
  \[ E_\mu [\tau_2(\mu)] \leq \sum_{i > N} \lambda_i^2. \]
- In the sequel we make this choice for \( \{ \phi_i \} \).
Bound – Estimation

In practice, we estimate

$$G\phi = \frac{1}{2} E_\mu \left( \langle r(\mu), \phi \rangle r(\mu) + \langle w(\mu), \phi \rangle w(\mu) \right).$$

by:

$$\hat{G}\phi = \frac{1}{2\#\Xi} \sum_{\mu \in \Xi} \left( \langle r(\mu), \phi \rangle r(\mu) + \langle w(\mu), \phi \rangle w(\mu) \right)$$

where $\Xi \subset \mathcal{P}$ is a sample of the distribution of $\mu$.

Matricially, the problem of finding $\phi_i$ is an eigenproblem in dimension $\min(N, 2\#\Xi)$. 
Bound – Majoration in probability

- We can estimate \( E_\mu [\tau_2(\mu)] \) by:

\[
\hat{T}_2 = \frac{1}{2\#\Xi} \sum_{\mu \in \Xi} \left| \bar{s}(\mu) - s(\mu) - \sum_{i=1}^{N} \langle w(\mu), \phi_i \rangle \langle r(\mu), \phi_i \rangle \right|,
\]

once for all the values of \( \mu \).

- Then, for a risk level \( \alpha \in ]0, 1[ \), we use Markov inequality:

\[
P_\mu(\tau_2(\mu) > E_\mu [\tau_2(\mu)] / \alpha) < \alpha,
\]

leading to an empirical threshold:

\[
\hat{T}_2 / \alpha.
\]

- And we have the final error bound estimate (with risk < \( \alpha \)):

\[
T_1(\mu) + \frac{\hat{T}_2}{\alpha},
\]

where (remember!) \( T_1(\mu) \) is a majorant of

\[
\left| \sum_{i=1}^{N} \langle r(\mu), \phi_i \rangle \langle w(\mu), \phi_i \rangle \right|.
\]
Correction of output

- The adjoint (dual) problem:

\[ A(\mu)^t w(\mu) = l, \]

can also be projected by using a matrix \( Z_d \):

\[ [Z^t_d A(\mu)^t Z_d] \tilde{w}(\mu) = Z^t_d l, \]

so as to given an approximation \( Z_d \tilde{w}(\mu) \approx w(\mu) \).

- Computation of \( \tilde{w}(\mu) \) generally doubles the computational time, but allows to compute a corrected output approximation for \( s(\mu) \):

\[ \tilde{s}_c(\mu) = \tilde{s}(\mu) - \langle Z_d \tilde{w}(\mu), r(\mu) \rangle, \]

which is known to be more precise than \( \tilde{s}(\mu) \).
More specifically, we can show that

$$|\tilde{s}_c(\mu) - s(\mu)| \leq \epsilon_u(\mu) \epsilon^d_u(\mu),$$

where $$\|w(\mu) - Z_d \tilde{w}(\mu)\| \leq \epsilon^d_u(\mu).$$

Our error bound can be readily extended so as to provide a bound $$\epsilon_c(\mu)$$ on the corrected output:

$$|\tilde{s}_c(\mu) - s(\mu)| \leq \epsilon_c(\mu),$$

in probability (with respect to $$\mu$$), by changing every $$w(\mu)$$ by $$w(\mu) - Z_d \tilde{w}(\mu).$$
There are four error bounds:

- on the non-corrected output:
  - Lipschitz bound: simple, deterministic but pessimistic;
  - our proposed bound on the non-corrected output: in probability, hopefully more accurate.

- on the corrected output (more expensive to compute, known to be more precise):
  - the existing bound in the literature;
  - our proposed bound on the corrected output.
Numerical result 1

Discretized PDE: diffusion.
Parametrisation of the geometry of the domain (3 parameters); risk $10^{-5}$.

![Graph showing numerical results](image)
Numerical result 2

Discretized PDE: transport (space-time formulation). 1 parameter (transport speed); risk $10^{-5}$. 

![Graph showing comparisons between different bounds and error estimates.](image-url)
Application to certified Sobol sensitivity analysis OK, thanks to the possibility of taking very small risks, avoiding the “multiple tests problem”.

Main perspective: application to non-linear models and/or non-linear outputs.
Michael Jackson’s
This Is It