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 μ.
- 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 $\mathcal{N} \gg 1$.
- 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 $\mathcal{N}$ lines, we look for $\widetilde{u}(\mu) \in \mathbf{R}^n$ so that:

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

- The system is of dimension n. Fine if $n \ll \mathcal{N}$.
- 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 $u(\mu) = Z\widetilde{u}(\mu)$.
- In many interesting cases, we have methods to choose Z so that

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

  and so:

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

- $\widetilde{s}(\mu)$: metamodel.
- Can we quantify the error in this approximation ?

Context (3) – Reduced basis error bound

- Under some hypotheses on the $A(\mu)$ matrix and a norm $\|\cdot\|$ (say, Euclidean norm), the reduced basis comes with an error bound $\epsilon^u(\mu)$:

$$\forall \mu \in \mathcal{P}, \|u(\mu) - Z\widetilde{u}(\mu)\| \leq \epsilon^u(\mu)$$

  which can be numerically computed *efficiently* (i.e., with the order of complexity of the computation of $\widetilde{u}(\mu)$).
- **Question:** Given this bound, can we have an error bound $\epsilon(\mu)$ on s:

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

  which can be explicitly and efficiently computed ?
- Yes, as the "Lipschitz bound" holds:

$$\forall \mu \in \mathcal{P}, |s(\mu) - \widetilde{s}(\mu)| \leq L\epsilon^u(\mu),$$

  for:

$$L = \sup_{\|v\|=1} l^t v.$$

Context (4) – Improved error bound

- **Question:** can we find a more precise error bound ?
- The Lipschitz bound is optimal amongst the bounds which depend on (a bound on) $\|u(\mu) - A(\mu)\widetilde{u}(\mu)\|$.
- 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

Reference: Janon, Nodet, Prieur, *Goal-oriented error estimation for reduced basis method, with application to certified sensitivity analysis*, submitted (HAL, arXiv).

Starting point

- Remember: $A(\mu)Z\widetilde{u}(\mu) \approx f(\mu)$.
- The bound $\epsilon^u(\mu)$ on $\|u(\mu) - \widetilde{u}(\mu)\|$ is based on the **residual:**

$$r(\mu) = A(\mu)Z\widetilde{u}(\mu) - f(\mu),$$

  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 ϕ are also efficiently computable.
- Let $\{\phi_i\}_{i=1,\dots,\mathcal{N}}$ be an orthonormal basis of $\mathbb{R}^{\mathcal{N}}$ (to be choosed later). We have:

$$\widetilde{s}(\mu) - s(\mu) = \sum_{i \geq 1} \langle w(\mu), \phi_i \rangle \langle r(\mu), \phi_i \rangle,$$

  where $w(\mu)$ is the solution of the **adjoint** (or **dual**) problem:

$$w(\mu) = A(\mu)^{-t} l,$$

  we set $\phi_i = 0$ for $i > \mathcal{N}$.

Error bound – Two-part decomposition

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

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

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

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

Bound – Addressment of the first term

- Let:

$$\tau_1(\mu) := \left| \sum_{i=1}^{N} \underbrace{\langle w(\mu), \phi_i \rangle}_{\textbf{to bound}} \overbrace{\langle r(\mu), \phi_i \rangle}^{\text{computable}} \right|$$

- We compute (once for all the values of μ):

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

where:

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

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

- We set:

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

and we have:

$$|\tau_1(\mu)| \leq \max \left(\left| \sum_{i=1}^{N} \langle r(\mu), \phi_i \rangle \beta_i^{low}(\mu) \right|, \left| \sum_{i=1}^{N} \langle r(\mu), \phi_i \rangle \beta_i^{up}(\mu) \right| \right) =: T_1(\mu).$$

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 μ is a random variable on $\mathcal{P}$, with known distribution.
- We want to control $\mathbf{E}_\mu \left[\tau_2(\mu) \right]$.
- We have:

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

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

$$\forall \phi \in X, \quad G\phi = \frac{1}{2} \mathbf{E}_\mu \left(\langle r(\mu), \phi \rangle r(\mu) + \langle w(\mu), \phi \rangle w(\mu) \right).$$

Bound – Addressment of the second term (2)

- Recall that:

$$\mathbf{E}_{\mu}\left[\tau_2(\mu)\right] \leq \sum_{i>N} \langle G\phi_i, \phi_i \rangle.$$

- Let $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_{\mathcal{N}} \geq 0$ be the eigenvalues of G, and ϕ_i^G a unitary eigenvector of G with respect to λ_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 τ_2:

$$\mathbf{E}_{\mu}\left[\tau_2(\mu)\right] \leq \sum_{i>N} \lambda_i^2.$$

- In the sequel we make this choice for $\{\phi_i\}$.

Bound – Estimation

- In practice, we estimate

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

  by:

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

  where $\Xi \subset \mathcal{P}$ is a sample of the distribution of μ.
- Matricially, the problem of finding ϕ_i is an eigenproblem in dimension $\min(\mathcal{N}, 2\#\Xi)$.

Bound – Majoration in probability

- We can estimate $\mathbf{E}_\mu\left[\tau_2(\mu)\right]$ by:

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

  once for all the values of μ.
- Then, for a risk level $\alpha \in]0,1[$, we use Markov inequality:

$$P_\mu(\tau_2(\mu) > \mathbf{E}_\mu\left[\tau_2(\mu)\right]/\alpha) < \alpha,$$

  leading to an empirical threshold:

$$\widehat{T}_2/\alpha.$$

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

$$T_1(\mu) + \frac{\widehat{T}_2}{\alpha},$$

  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:

$$[Z_d^t A(\mu)^t Z_d]\widetilde{w}(\mu) = Z_d^t l,$$

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

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

$$\widetilde{s}_c(\mu) = \widetilde{s}(\mu) - \langle Z_d \widetilde{w}(\mu), r(\mu)\rangle,$$

which is known to be more precise than $\widetilde{s}(\mu)$.

Correction of output (2)

- More specifically, we can show that

$$|\widetilde{s}_c(\mu) - s(\mu)| \leq \epsilon_u(\mu)\epsilon_u^d(\mu),$$

where $\|w(\mu) - Z_d\widetilde{w}(\mu)\| \leq \epsilon_u^d(\mu)$.

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

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

in probability (with respect to μ), by changing every $w(\mu)$ by $w(\mu) - Z_d\widetilde{w}(\mu)$.

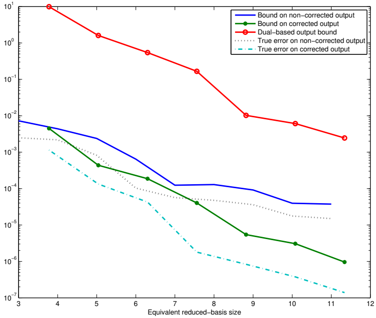
Summary

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.

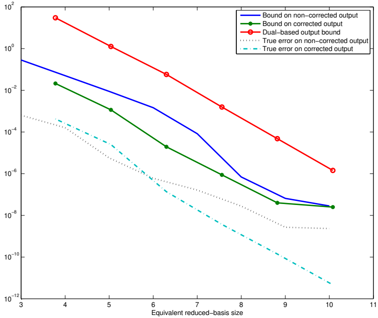
Numerical result 1

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



Numerical result 2

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



Concluding remarks and perspectives

- 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