

PRINCIPAL ANGLES BETWEEN SUBSPACES AND REDUCED ORDER MODELLING ACCURACY IN SCALABLE ROBUST OPTIMIZATION

AN EESI2 CONTRIBUTION TO THE EMWG WORKSHOP

BIJAN MOHAMMADI
UNIVERSITÉ MONTPELLIER II, MATHÉMATIQUES
CC51, 34095 MONTPELLIER, FRANCE

ABSTRACT. The paper considers robust parametric optimization problems using multi-point formulations and addresses the issue of the approximation of the gradient of the functional by reduced order models. The question of interest is the impact of such approximations on the search subspace in the multi-point optimization problem. The mathematical concept used to evaluate these approximations is the principal angles between subspaces and practical ways to evaluate these are provided. An additional indicator is provided when a descent minimization algorithm is used. The approach appears also to be an interesting tool for uncertainty quantification in the presence of models of increasing complexity. We are especially interested by the choices which bring into light efficient parallelism. The application of these concepts is illustrated in the design of the shape of an aircraft robust over a range of transverse winds.

1. EMWG/EESI2 CONTEXT

The following considerations motivate this work from the point of view of calculation complexity:

- today the solution of the state equations of interest (e.g. from fluid mechanics or seismic) already mobilizes petaflop calculation facilities available. This is called the solution of the direct problem.
- the best calculation complexity one might think of for a robust parameteric optimization is when its cost is made comparable to a mono-point situation.
- exascale supercomputers open new horizons for large simulations, but one should still think of calculation efficiency, because even those will not permit the solution of robust design problems where the direct state requires a petaflop computer.

We will show here that not only brute force (e.g. by extensive sampling) is not an answer, but it brings to these robust design problems misleading informations with noise and numerical artifacts encapsulated.

This work is an example of how to find new formulations and algorithms in order to take advantage of the next generation of supercomputers in order to keep the calculation complexity optimal (i.e. of a mono-point design).

2. INTRODUCTION

Reduced order models are widely used in both simulation and design. These can be built, for instance, by assimilation of high fidelity data or experimental results by a parametric model. Once the model built it can be used in an optimization procedure providing cheap estimation of the functional and its sensitivity with respect to the model parameters. It

Key words and phrases. robust optimization, parallelism, reduced order modelling, principal angles between subspaces, uncertainty quantification.

can also be locally used in trust region approaches where the reduced order model needs to be dynamically enriched. Reduced order modelling receives different denominations following the field of research it is issued from: learning in neural networks, fitting with least squares in experimental plans or higher degrees polynomials, reduced-order modeling with proper orthogonal decomposition, etc. [1, 2, 3].

The performance of a system designed for given functioning conditions often seriously degrades when these conditions are modified. Typical situations of interest involve a few (typically one or two) parameters describing the functioning of the system. Multi-point optimization is widely used to address robustness issues in engineering and we showed [17, 12] how to use such formulation to address the robustness issue when a few functioning parameters are involved. In particular, we showed how to introduce in this formulation what we would like for the outcome of the design through a 'target-based' weighting in the functional. We also discussed optimal sampling of the functioning parameters intervals and quantitative confidence levels on the quality of our search direction through Gram-Schmidt orthonormalization of the multi-point gradients. This naturally brought into light the question of the sampling size which needs to be monitored to maintain the calculation complexity low. The aim was to avoid the worst-case theoretical limitation indicating that the sampling size should be larger by one the size of the control space [19, 17, 18]. We showed that this is only necessary if all the associated gradients (i.e. evaluated at the sampling points) are linearly independent which is never the case in optimizations involving a state equation. Hence, we showed, again through Gram-Schmidt procedure, that large dimensional parametric optimization problems can be treated with very small sampling of the functioning parameters range with marginal losses on the gradient informations.

However, one often uses approximation in the definition of the gradients. And this is the starting point of this paper. One would like to analyze the impact of these using reduced order models on the search space defined by the gradients of the functional at the different sampling points of the functioning parameters. The mathematical tool we use for this analysis is the principal angles between subspaces. Reduced order modelling also includes situations where lower accuracy and modelling complexity are accounted for in the linearization step than for the direct simulation chain used for the definition of the functional. This is often the situation where an approximate simulation chain is considered for linearization, dropping or approximating some of the ingredients in the direct chain.

3. ROBUST PARAMETRIC OPTIMIZATION

4. A MULTI-POINT DESCENT ALGORITHM

We consider the following iterative descent algorithm for our constrained minimization problem involving a direct simulation chain linking the parameters (x, α) to the state U solution of a state equation $F(U(q(x, \alpha))) = 0$ and to a functional j is:

$$\left\{ \begin{array}{l} \text{Given } x_0, 0 < \rho, 0 < \eta \ll 1, \mathbf{I}_m, p_{max}, 0 < \varepsilon \ll 1, \\ \text{optimization iterations } p = 1, \dots, p_{max} \\ \left\{ \begin{array}{l} -m \text{ parallel state equation solutions } F(U(q(x_p), \alpha_l)) = 0, \alpha_l \in \mathbf{I}_m, \\ -m \text{ parallel evaluations of } j(x_p, \alpha_l), \alpha_l \in \mathbf{I}_m, \\ -m \text{ parallel solutions of the adjoint state } V \text{ equation:} \\ \qquad V^t F_U(U(q(x_p), \alpha_l)) = j_U^t, \alpha_l \in \mathbf{I}_m, \\ -m \text{ parallel evaluations of } \nabla_x j(x_p, \alpha_l) = j_x + (V^t F_x)^t, \alpha_l \in \mathbf{I}_m, \\ -\text{define } d \text{ the descent direction, } d = \overline{\nabla_x \mu} - \langle \overline{\nabla_x \mu}, \overline{\nabla_x \sigma} \rangle \overline{\nabla_x \sigma} + \eta \overline{\nabla_x \sigma}, \\ -\text{control parameter variation, } x_{p+1} = x_p - \rho d, \\ \text{Stop if } \|d\| \leq \varepsilon, \end{array} \right. \end{array} \right.$$

where $\bar{a} = a/\|a\|$ is the normalized vector a and

$$(1) \quad \nabla_x \mu = \frac{1}{\Omega} \sum_{\alpha_l \in \mathbf{I}_m} \omega_l \nabla_x j(x, \alpha_l),$$

$$(2) \quad \nabla_x \sigma = \sum_{\alpha_l \in \mathbf{I}_m} \langle \nabla_{\alpha} j(x, \alpha_l), \nabla_{\alpha x} j(x, \alpha_l) \rangle .$$

$\nabla_{x\alpha} j(x, \alpha_l)$ is obtained from $\nabla_x j(x, \alpha_l)$ by finite differences on \mathbf{I}_m , component by component, following what has been done for $\nabla_{\alpha} j$. The definition of the descent direction permits to make sure that both μ and σ decrease for small descent steps. Indeed, a first order development in x gives:

$$\sigma(x_{p+1}) - \sigma(x_p) = \|\nabla_x \sigma\| \overline{\nabla_x \sigma} \cdot (x_{p+1} - x_p) = -\rho \eta \|\nabla_x \sigma\| \leq 0,$$

and we have

$$\mu(x_{p+1}) - \mu(x_p) = \|\nabla_x \mu\| \overline{\nabla_x \mu} \cdot (x_{p+1} - x_p) = -\rho \|\nabla_x \mu\| (1 - \zeta^2 + \eta \zeta),$$

where $\zeta = \langle \overline{\nabla_x \mu}, \overline{\nabla_x \sigma} \rangle$. Therefore, μ is also decreasing as $1 - \zeta^2 + \eta \zeta \geq 0$ for $|\zeta| \leq 1$ and $0 < \eta < 1$ as chosen in the algorithm.

Parallelism: Despite the natural presence of parallelism in this algorithm in the m independent evaluations of the state, functional and its gradient, computational complexity remains an issue. Possible solutions for the reduction of the sampling size have been presented in [12] together with the use of incomplete sensitivity concept in the evaluation of the gradients. This latter permits to avoid the solution of the m adjoint equations in the algorithm. This is particularly suitable when using black-box state equation solvers not providing the adjoint of the state variables.

Beyond individual gradient accuracy (i.e. at each of the sampling point), what is important in these multi-point problems is the global search space defined by the ensemble of the gradient vectors. This means that one might tolerate higher error levels in each of the gradient defined at the different sampling point than for a mono-point optimization situation as what is important is for the global search space to remain nearly unchanged. An interesting mathematical concept which permits to measure the deviation between two subspaces is the principal angles between subspaces.

5. ANGLES BETWEEN SUBSPACES

6. DEFINITION OF THE SUBSPACES

7. EVOLUTION OF THE PRINCIPAL ANGLES DURING OPTIMIZATION

8. SENSITIVITY ANALYSIS

8.1. Examples of reduced order models.

8.1.1. *Model reduction on the continuous level.*

8.1.2. *Complexity reduction on the discrete level.*

8.1.3. *Evaluation of an incremental modelling process.*

8.2. Defining the descent direction.

9. FULL AIRCRAFT SHAPE OPTIMIZATION

REFERENCES

- [1] Schilders, W. H. Van der Vorst, H. A. Rommes, J. *Model order reduction: Theory, research aspects and applications*, Springer Math in Industry series Vol. 13, Berlin, (2008).
- [2] Obinata, G. Andersonn B. *Model reduction for control system design*, Springer, Berlin (2000).
- [3] Qu, Z. *Model Order Reduction Techniques with Applications in Finite Element Analysis*, Springer, Berlin (2004).
- [4] Jordan, C. Essay on geometry in n dimensions, *Bull. Soc. Math. France*, **3**, 103:174 (1875).
- [5] Gluck, H. Warner, F. Great circle fibrations of the three-sphere, *Duke Math. J.*, **50**, 107-132 (1983).
- [6] Morano, E. Mavriplis, D. Venkatakrisnan, V. Coarsening strategies for unstructured multigrid techniques with application to anisotropic problems, *SIAM J. Sci. Comput.*, **20-2**, 393:415 (1998).
- [7] Lallemand, M.-H. Steve, H. Dervieux, A. Unstructured multigriding by volume agglomeration: Current status, *Computers and Fluids*, **21**, 397:433 (1992).
- [8] Jiang, S. Angles between Euclidean subspaces, *Geometricae Dedicata*, **63(2)**, 113-121 (1996).
- [9] Shonkwiler, C. Poincare duality angles for Riemannian manifolds with boundary, PhD. thesis, Univ. Pennsylvania (2009).
- [10] Mohammadi B., Pironneau O. *Applied Shape Optimization for Fluids* (2nd Edition), Oxford Univ. Press, London (2009).
- [11] Griewank, A. *Computational derivatives*, Springer, New York, (2001).
- [12] Mohammadi B. Reduced sampling and incomplete sensitivity for low-complexity robust parametric optimization, *Int. J. Num. Meth. Fluids*, DOI: 10.1002/fld.3798 (2013).
- [13] Mohammadi B. Value at Risk for confidence level quantifications in robust engineering optimization, *optimal Control: Applications and Methods*, DOI: 10.1002/oca.2061 (2013).
- [14] Melchers, R. E. *Structural Reliability Analysis and Prediction*. John Wiley and Sons, Chichester (1999).
- [15] Mohammadi, B. Pironneau, O. Shape Optimization in Fluid Mechanics, *Annual Revue of Fluid Mechanics*, **36/1**, 255:279 (2004).
- [16] Hascoet, L. Pascual, V. Tapenade 2.1 user's guide. INRIA Technical Report RT-300 (2004).
- [17] Gallard, F. Mohammadi, B. Montagnac, M. Meaux, M. Robust parametric design by multi-point optimization. CERFACS Technical Report TR/CFD/12/33 (2012).
- [18] Desideri, A. Multiple-gradient descent algorithm for multiobjective optimization. CRAS. **350/5**, 313:318 (2012).
- [19] Li, W. Huyse, L. and Padula, S. Robust Airfoil Optimization to Achieve Consistent Drag Reduction Over a Mach Range, *Structural and Multidisciplinary Optimization*, **24/1**, 38-50 (2002).