Optimization without derivatives under constraints : with surrogate quadratic models in a trust region

Hoël Langouët, PhD student of IFP, Rueil-Malmaison, France, hoel.langouet@ifp.fr

---

Optimization takes place in many IFP applications: estimation of parameters of numerical models from experimental data (earth sciences, combustion in engines), design optimization (networks of oil pipelines), optimizing settings of experimental devices (engine calibration, catalysis). These optimization problems consist in minimizing a functional that is complex (nonlinearities, noise) and expensive to estimate (solution of a numerical model based on differential systems or experimental measurements), and for which derivatives are often not available, with nonlinear constraints, and sometimes with several objectives in order to find the best compromise.

At first, we are interested in minimizing, under linear constraints, a function, which is expensive to evaluate, and for which the derivatives are not available:

$$\begin{cases} \min_{x \in \mathbb{R}^n} f(x), \\ b_l \leqslant x \leqslant b_u, \quad b_l, b_u \in \mathbb{R}^n, \\ Ax \leqslant b, \qquad A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m. \end{cases}$$

In practice, these problems are often resolved by non linear optimization methods subject to constraints (SQP [10] for example) with derivatives approximated by finite differences. Although these methods are particularly efficient for the determination of active constraints, the cost in evaluations of the function to be minimized is usually too high for industrial problems with expensive simulators. Furthermore, the choice of step for the finite differences, crucial for the convergence of this method, is generally very difficult, because it depends on the accuracy of f which is difficult to estimate in practice.

Direct Methods optimize a function without derivative calculation. Among them, Genetic Algorithms [2] and Pattern Search methods [4] are insensitive to inaccuracies in the calculation of f but require many evaluations. To overcome this difficulty, a cheap surrogate model of the objective function is typically used in the optimization to limit the number of evaluations of the expensive function. These surrogate models are usually global models of f on the domain under study, constructed from a limited number of evaluations of f chosen according to a relevant criteria related to the estimation of the prediction error of these models and expected improvement on the minimization of the function [3, 12]. These models can be kriging models [3, 9, 12], Radial Basis Functions (RBF), splines ... But they are limited to problems of small size (10-20 param.).

Then another class of methods based on local surrogate models was proposed by [1, 7, 8, 11]: these methods are inspired by SQP methods with trust region globalization [1]. A quadratic model is built at each iteration in a neighborhood of the current point, the size of this neighborhood are updated according to the comparison of the reduction predicted by the model and the effective reduction calculated by evaluating f. In particular, Powell [7, 8] proposed a very efficient method, without constraints on problems of medium size (100 param.).

Taking into account constraints in these methods remains a difficulty, penalty methods are often inefficient in practice. In this paper, an extension of Powell's method to linear constraints is proposed, it consists in :

1. choosing the interpolation points in the domain defined by the constraints,
2. minimizing the quadratic model under constraints in the trust region thanks to a SQP method.

Numerical results for noisy test functions (relative noise level of 10^{-3}) constructed from the benchmark CUTEr [5] are presented. Bound constraints on parameters are introduced. The results obtained with the SQA method are compared with two methods :

1. SQPAL, Sequential Quadratic method with gradient of the Lagrangian estimated by finite differences (step of 10^{-3} and 10^{-6}) and a BFGS approximation of the Hessian matrix,
2. EGO, a sequential kriging method with the addition of points using expected improvement criteria [9].

Data profiles and performance profiles [5] are used to compare the three methods. They are presented for two accuracies degined by : $\tau = \frac{f(x^*)-f_L}{f(x_0)-f_L}$.

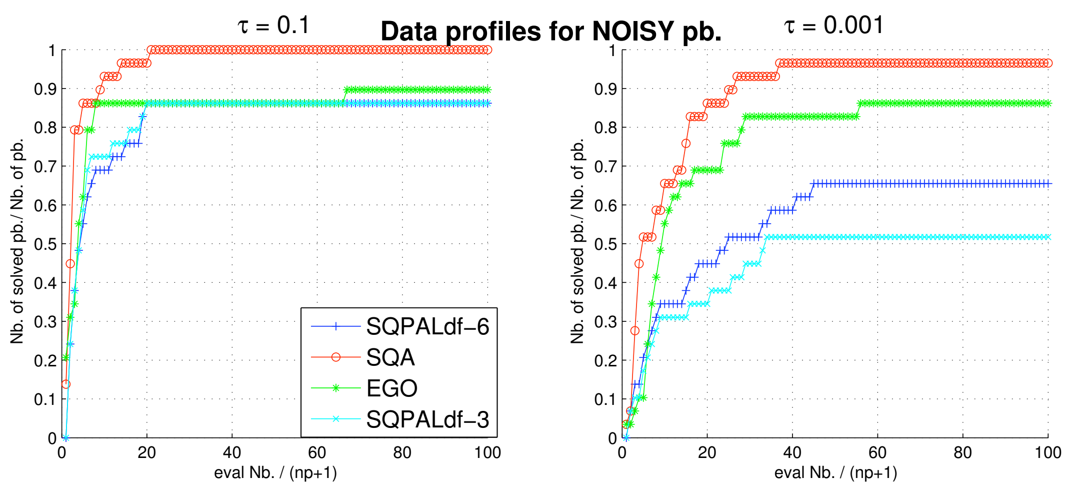

Figure 1: Data profiles for noisy problems

Data profiles in Figure 1 show that the method SQA solves the largest percentage of problems for all function evaluation budgets and for all required accuracy τ.

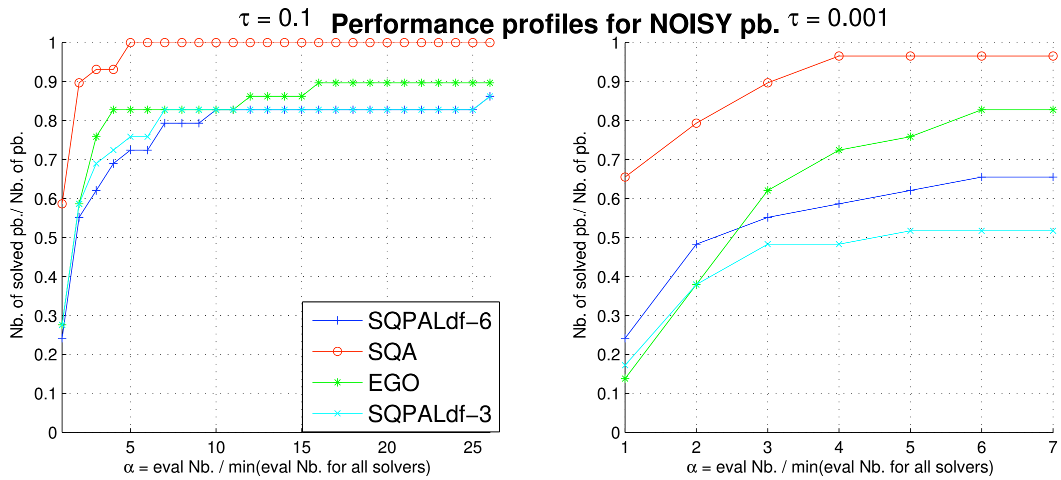

Figure 2: Performance profiles for noisy problems

Performance profiles of Figure 2 show that the method SQA is the fastest solver in at least 60% of the problems, while EGO and SQPAL are each the faster for less than 30% of the problems ($\alpha = 1$). In addition, the SQA method is the most robust because it solves the largest percentage of problems for all levels of accuracy τ ($\alpha \to \infty$).

An application for production and seismic data matching in the field of oil reservoir characterization will also be presented.

References

[1] Conn A. R., Gould N. I. M. et Toint P. L., (2000) *Trust-Region Methods*, MPS-SIAM Series on Optimization.

[2] Hansen, N. et Kern N., (2004) *Evaluating the CMA Evolution Strategy on Multimodal Test Functions*, In Eighth International Conference on Parallel Problem Solving from Nature PPSN VIII, Proceedings, pp. 282-291, Berlin: Springer.

[3] Jones D., Schonlau M. et Welch W., (1998) *Efficient global optimization of expensive black-box functions.*

[4] Kolda T. G., Lewis R; M. et Torczon V., (2003) *Optimization by Direct Search: New Perspectives on Some Classical and Modern Methods*, SIAM Review, Vol. 45, No. 3, pp. 385-482.

[5] Moré J. J. et Wild S. M., (2007) *Benchmarking Derivative-Free Optimization Algorithms*, Mathematics and Computer Science Division, Argonne National Laboratory, Preprint ANL/MCS-P1471-1207, http://www.mcs.anl.gov/ more/dfo/ceperf.pdf.

[6] Oeuvray R., (2005) *Trust-Region Methods Based on Radial Basis Functions with application to Biomedical Imaging*, Thèse de doctorat, Ecole Polytechnique de Lausanne.

[7] Powell M.J.D., (2004) *The NEWUOA software for unconstrained optimization without derivatives.*

[8] Powell M.J.D., (2007) *Developments of NEWUOA for unconstrained minimization without derivatives.*

[9] Schonlau M., (1997) *Computer experiments and global optimization*, Thèse de doctorat, Université de Waterloo.

[10] Sinoquet D. et Delbos F., (2008) *Adapted nonlinear optimization method for production data and 4D seismic inversion*, 11th ECMOR, Bergen, Norway, 8 - 11 September 2008.

[11] Vanden Berghen F., (2004) *CONDOR : a constrained, non-linear, derivative-free parallel optimizer for continuous, high computing load, noisy objective functions*, Thèse de doctorat, Université de Bruxelles.

[12] Villemonteix J., (2008) *Optimisation de fonctions coûteuses, Modèles gaussiens pour une utilisation efficace du budget d'évaluations : théorie et pratique industrielle*, Thèse de doctorat, Université Paris-sud XI.

[13] Wild S. M., Regis R. G. et Shoemaker C. A., (2008) *ORBIT: Optimization by Radial Basis Function Interpolation in Trust-Regions*, SIAM J. Scientific Computing, Vol. 30, No. 6, pp. 3197-3219.