Derivative free optimization: without and with constraints

Philippe L. Toint

University of Namur, Belgium

ABSTRACT

The problem is that of computing the solution of nonlinear programming problem for which the value(s) of the objective function (and, if any, of the constraints) is available, but without any possibility to compute the derivatives of these functions. Such situations occur in many practical industrial environments. For instance, it may be that the objective function value is derived from the actual measurement of a physical quantity, or that its value results from an adaptive discretization of a PDE problem, or, most commonly, that the value is derived from a complicated computer-based simulation, for which the source code is unavailable. Furthermore, it often happens that the objective function value is also contaminated by noise, which obviously creates other problems.

The class of methods that will be discussed in the talk is based on the idea that the typical very expensive function values could be used in a more sophisticated way than for computing finite-difference approximations to derivatives (which would, anyway, be irrelevant for noisy problems). Instead the main idea is to use the techniques of multivariate interpolation to construct appropriate models of the objective function, and to use them in the context of a more or less classical trust-region minimization algorithm. Unfortunately, the interpolation conditions alone are not enough to guarantee that the resulting algorithm is well-defined. It is also vital that the model be "valid"' in the sense that certain geometrical conditions, known as poisedness, must also hold before a model is deemed inadequate by the algorithm. But these conditions may be checked during the construction of the interpolation model itself, which makes the construction of a suitable method possible. One such approach will be described where the interpolation models are based on the "Newton fundamental polynomials" and associated "generalized finite differences". The nature of these polynomials and their role in the interpolation procedure will be described.

Using these tools, a trust-region algorithm for unconstrained minimization will be introduced, with a brief outline of the trust-region mechanism, an essential ingredient for ensuring robustness of the method. Some numerical results will be presented that show the potential of the method. If time allows, a glimpse of the convergence theory will also be presented.

The discussion will be continued by addressing ways in which known structure of the minimization problem can be taken into account. In particular, the use of partial polynomial basis for taking Hessian sparsity into account will be reported. This covers the necessary adaptations of the minimization algorithm and additional numerical tests.

In the same vein, we will next concentrate on how structure of the problem might be discovered, which will lead to the introduction of partially separable problems, a concept inherited from large-scale quasi-Newton methods. Further adaptations of the algorithm to

this case will be outlined, and the three approaches (unstructured, sparse and partially separable) will be contrasted.

In a second part of this talk, we will consider adaptations of the same ideas to the case where constraints are also present. The approach used will be that of “filter methods", a new class of constrained optimization methods that is the subject of active ongoing research. First the concept of a filter method will be introduced, in the general case where derivatives are available. The basic nature of the filter will be described and the organization of the resulting algorithm outlined. The resulting method is of the Sequential Quadratic Programming (SQP) type, probably the most powerful class of optimization methods known today. The basic framework of SQP methods will be recalled.

The discussion will be concluded with the description of how the presented techniques can be used in conjunction with interpolation models, and some numerical results again presented.

<u>Note:</u>

A. Conn, K. Scheinberg and B. Colson are co-authors of this paper.