Abstract

Most of the optimization algorithms depend on derivative information of the problem. However, there are numerous real-world problems where derivatives are unavailable. This motivates us to study derivative-free optimization methods.

The assessment and comparison of algorithms play important roles in the research of algorithms. We study how to assess derivative-free algorithms in a reliable way. Through an example, we show that it is not reliable to merely count the number of function evaluations. By introducing statistical method, we establish a new system for the assessment of derivative-free algorithms. The new system reflects the stability of algorithms with respect to computer rounding errors, and provides more convincing comparison of different algorithms.

Least Frobenius norm quadratic interpolation and symmetric Broyden update are the most successful methods of constructing models in derivative-free trust-region algorithms. We prove the equivalence between these two strategies in some cases. The restart technique in `NEWUOA` is closely related to these two strategies. We modify the restart criterion in the source code of `NEWUOA` by simply deleting four letters and obtain a new version of the source code. The modification brings considerable improvement. Under the framework of `NEWUOA`, we compare least Frobenius norm model and the model established by symmetric Broyden update. We point out that least Frobenius norm model works better if a low-precision minimizer is desirable, which is meaningful for applications, because many problems in practice do not require high-precision solutions.

To study the widely-used least norm quadratic interpolation in derivative-free methods, we introduce the Sobolev norms and seminorms, which are classical in PDE theory but rarely noticed in optimization. For the H^0 norm and H^1 seminorm of a quadratic function over an ℓ_p ball, we obtain explicit formulae in terms of the coefficients of the function. We prove that least norm quadratic interpolation seeks an interpolant with minimal H^1 seminorm over an ℓ_2 ball. This observation provides a new perspective to interpret the interpolation. Consequently, we present the geometrical meaning of the parameters in the interpolation. We apply our theory to study the extended symmetric Broyden update, and propose a very simple but effective way of choosing the parameters in the update.

Until now, derivative-free methods can only solve problems with modest dimension. We study subspace techniques to attack large scale problems. We presents two derivative-free subspace methods. In the first method, we apply the idea of Hooke-Jeeves pattern-search to derivative-free trust-region method, and propose to solve the trust-region subproblem in a low-dimensional subspace of $\mathbb{R}^n$. This subspace strategy improves the performance of `NEWUOA`. The second subspace method, which is named as `NEWUOAs`, is the highlight of the thesis. Its basic idea is to divide a large scale problem into a sequence of low-dimensional subproblems. We first study a general subspace algorithm based on this idea, and present its convergence theory. Then using `NEWUOA` as the subproblem solver, we implement the algorithm without using derivatives and obtain `NEWUOAs`. For `NEWUOAs`, we establish its global convergence and R-linear convergence rate in theory, and prove its finite termination in computation. We propose a preconditioning technique, which improves the performance of `NEWUOAs` on ill-conditioned problems. As far as we know, this is the first derivative-free algorithm with preconditioning procedure. In our numerical experiments, `NEWUOAs` works evidently better than `NEWUOA`, in the number of function evaluations, CPU time, and stability. We also find that `NEWUOAs` is good at solving problems with bad starting points, which is favourable in real-world applications. Besides, `NEWUOAs` is capable of solving many 2000-dimensional test problems to high precision within several minutes, using not more than 50000 function evaluations (equivalent to less than 25 simplex gradients). It is a breakthrough, because most state-of-the-art derivative-free algorithms can only solve problems with not more than a few hundreds of variables, and 2000-dimensional problems are nearly unsolvable for them.

Keywords: derivative-free optimization, trust-region method, quadratic interpolation, symmetric Broyden update, Sobolev seminorm, subspace method, large scale problem