Derivative-free Optimization with Combinatorial Properties

Juan J Torres[1], Emiliano Traversi[1], Roberto W Calvo[1] , Giacomo Nannicini[2]

[1] Université Paris 13, Sorbonne Paris Cité, LIPN, CNRS, (UMR 7030), France

{torresfigueroa,traversi,wolfer}@lipn.univ-paris13.fr

[2] IBM T.J. Watson Research Center, NY, U.S.A.

nannicini@us.ibm.com

Mots-clés : *Derivative-free optimization, Mixed-integer programming, Discrete convex functions, Surrogate-based optimization*

1 Introduction

In this project we propose an algorithm for the solution of the following problem :

$$\min_{x,y} f(x,y) \tag{1a}$$

$$x_{lb} \leq x \leq x_{ub}, \quad y_{lb} \leq y \leq y_{ub} \tag{1b}$$

$$x \in \mathbb{R}^{n_1}, \quad y \in \mathbb{Z}^{n_2} \tag{1c}$$

where $f : [\mathbb{R}^{n_1} \times \mathbb{Z}^{n_2}] \to \mathbb{R}$ is a mixed-integer black-box function which exhibits combinatorial properties at fixed values of x. Black-box are often expensive-to-evaluate and do not present analytical form, which means that no gradient nor second-order information can be used to optimize them. Black-box functions arise in several settings, such medical imaging, operations research, and, specially in computer simulation programs. Several methodologies have been developed to solve black-box instances including heuristics (i.e evolutionary algorithms, tabu search) and derivative-free optimization (DFO) methods. Within DFOs, surrogate-based approaches appear to be successful in the computation of local and global solution of black-box optimization programs.

Surrogate-based methods consist in the computation of a (surrogate) model which approximates the black-box function, via regression or interpolation. Different types of models can be use to approximate a black-box function, including low-order polynomials, radial basis functions (RBF) and kriging. To the extent of our knowledge, none of these approximations have been studied for mixed-integer functions with special combinatorial properties on their integer elements.

One example of such problems is the mixed-integer generalization of $M^\natural$ discrete functions [1]. $M^\natural$ are integrally convex functions that display interesting properties such as supermodularity, descent directions and minimizers. We aim to develop interpolants which can mimic this behavior, in order to improve the approximation of structured functions, and reduce the number of iterations on the optimization of black-box problems.

2 Algorithm Structure

Our proposed algorithm is designed to solve problem (1) via surrogate optimization. This is done by a three phase methodology which aims to approximate the mixed-integer function $f(x,y)$ with a suitable quadratic interpolant Q_k, which exhibits combinatorial properties. **Phase 1** consist on the initial sampling, **Phase 2** considers the computation of the surrogate model and geometry handling, and **Phase 3** is the computation of a new candidate solution via a special procedure derived from the difference of convex algorithm (DCA) [2]. Phase 2 and Phase 3 are repeated until algorithmic convergence is detected.

2.1 Phase 1 : Initial sampling

To compute an initial interpolant a design of experiments is performed, consisting in the evaluation of $f(x, y)$ on at least $2n + 1$ independent points, where n is the number of variables in the system. To assure a proper spanning of $\mathbf{dom} f$ we use a "Latin Hypercube design" with emphasis on a maximum/minimum correlation of samples.

2.2 Phase 2 : Model approximation

The quadratic approximation of $f(x, y)$ is the following :

$$Q_k(x, y) = c + g_C^T x + g_I^T y + \frac{1}{2} x^T A_C x + \frac{1}{2} y^T A_I y + y^T A_M x \tag{2}$$

where c is a constant term, g_C and g_I are linear coefficients, the interaction matrix A_M and the symmetric matrices A_I and A_C. All this coefficients are computed via regression using an adequate disjunctive formulation, devised to provide the matrix A_I with a combinatorial structure. To improve the quality of the approximation, a geometry handling procedure is activated when the number of sampling points reach a threshold. This procedure retires elements from regression based on criteria such value of the sampled objective or distance to the current best solution.

2.3 Phase 3 : Candidate computation

A new candidate is computed by solving the surrogate model inside a trust-region, centered in the best known solution. The chosen trust-region consider independent box-constrains, specially tailored for a mixed-integer domain. The trust-region is continuously updated based on the balance improvement of the objective function, and the quality of the regression. The algorithmic termination is reached when the trust-region radius is smaller than a user provided tolerance ϵ.

The solution procedure exploits the combinatorial properties of matrix A_I and performs a separation of integer and continuous variables. This methodology is well suited for cases when the relaxation of Q_k is convex and can be adapted via regularization to solve non-convex instances. Although it was originally developed as an local optimization technique, it provides high quality solutions when used as part of a multistart procedure. Numerical tests on quadratic mixed-integer $M^\natural$ instances show this methodology is faster to compute feasible, even optimal, solutions compared to available commercial solvers.

Références

[1] K. Murota, *Discrete convex analysis.* SIAM, 2003.

[2] P. D. Tao *et al.*, “The dc (difference of convex functions) programming and dca revisited with dc models of real world nonconvex optimization problems,” *Annals of operations research*, vol. 133, no. 1-4, pp. 23–46, 2005.