ON STATISTICAL MODELS BASED MULTI-OBJECTIVE OPTIMIZATION

A. ŽILINSKAS

*Institute of Mathematics and Informatics, Vilnius University,
Akademijos 4, LT-08663 Vilnius, Lithuania
e-mail: antanas.zilinskas@mii.vu.lt*

Abstract

Optimization problems with “expensive” and “black box” objectives are difficult to tackle. Some experience is accumulated in single-objective global optimization of those problems using algorithms based on statistical models of objective functions. The generalization of this approach to the multi-objective optimization is discussed.

1 Introduction

Nonlinear multi-objective optimization is very active research area. Depending on the properties of a multi-objective optimization problem, different approaches to its solution can be applied. The best direction developed is optimization of convex problems; for the latter problems the methods that generalize the ideas of classical mathematical programming suit well [2]. For the problems with not so nice objectives, metaheuristic methods are frequently favorable [1], [4]. However, there remains a class of important problems without sufficient attention of researchers: namely, the problems with “black box”, multimodal, and “expensive” objectives. A well established approach to single objective optimization of “black box”, multimodal, and “expensive” objectives is based on statical models of objectives. In the present paper we discuss the possibilities to generalize that approach to the multi-objective case. Global optimization (GO) of non-convex functions is a challenging problem. The development of single-objective GO algorithms for some sub-classes of non-convex functions is facilitated by the exploitation of analytical properties of the objectives. However, in some applications there occur optimization problems where objectives are available either as a computational model or as software. We focus on the problems where objective functions are “expensive” because of the complexity of the computational model; “expensiveness” here means long lasting computation of a value of the objective function. The complexity of the computational model normally implies not only the “expensiveness” of the objective function but also the uncertainty in its properties. “Black box” optimization of “expensive” functions with many respects is quite opposite to the optimization of objective functions defined by analytical formulae. The limitation in collecting general information about the function, and particularly about its minima, strongly requires the rationality in distribution the points where to compute the objective function values. Therefore the algorithms justified by the principles of rational decision theory here are of especial interest. To construct such algorithms in the single objective optimization case the statistical models of multimodal functions have been proven very helpful [3],

[5], [6]. Recently several papers have been published which propose multi-objective optimization algorithms generalizing single-objective optimization algorithms based on statistical models of objective functions. The proposed algorithms are straightforward generalizations of the single-objective prototypes, and their theoretical analysis is absent. In the present paper we discuss theoretical results justifying the use of statistical models for GO, and some properties of the corresponding single-objective optimization algorithms, aiming to facilitate the generalization of the GO approach based on statistical models to the multi-objective case.

2 Statistical models

Let consider the global minimization problem $\min_{X \in \mathbf{A}} f(X)$, $\mathbf{A} \subset R^d$, where $f(\cdot)$ is a continuous function and A is a compact set. The comments below specify the concept of “black box” optimization of “expensive” functions. It is assumed that $f(X)$ values can be either computed by a software implementing an unknown algorithm, or evaluated experimentally. Besides of continuity, other analytical properties of $f(X)$ are not known. It is supposed that the either long lasting computations or the experimentations are needed to evaluate a single value of $f(X)$. Such unfavorable from the optimization point of view properties as non-differentiability, non-convexity, and multimodality can not be excluded. To justify a search strategy in the described situation of uncertainty a “rational optimizer” should define a model of uncertainty. The general theory of rational decision making substantiates the use of statistical models of uncertainty. The first (one-dimensional) statistical model used for global black box optimization was Wiener process, i.e. it was assumed that an objective function to be optimized is a random sampling function of the Wiener process; see, e.g. [6]. In discussions, which followed the suggestion to use the Wiener process as a statistical model for GO, frequently it was counter argued that the local properties of sampling functions of the Wiener process are very different from those of the functions which are typical in applications. However, the discrepancy in local properties is not relevant to GO, as it was shown in [7], where also even more general statistical models have been justified. The most general statistical model axiomatically justified in [7] is a family of random variables $\xi(X)$ which can be interpreted very simply: a unknown function value of the objective function optimized is interpreted as a random variable. In the case of multi-objective optimization a vector objective function $F(X) = (f_1(X), f_2(x), \dots, f_m(X))^T$ is considered. The same arguments as in the case of single-objective optimization are applicable to prove that the compatibility of the attributes “black box” and “expensive” with the standard assumptions of rational decision making implies the acceptability of the family of random vectors $\Xi(X) = (\xi_1(X), \dots, \xi_m(X))^T$, $X \in \mathbf{A}$, as a statistical model of $F(X)$. Particularly, the vector of Wiener processes is a generalization of the frequently used one-dimensional statistical model to the multi-objective case. Similarly the vector valued random field $\Xi(X)$ is a generalization of statistical models of single objective GO where the components of $\Xi(X)$ are defined in the same way as single objective statistical models, e.g. homogeneous isotropic Gaussian random fields.

3 Multi-objective algorithms based on statistical models

he rationality principles with respect to a search method informally can be formulated as suggestions related to the choice of the current point for computation of the vector of objectives:

- it is not rational to choose the current point in the vicinities of points where function values have been computed at previous steps since it would give scarce global information about objectives,
- the points, where expected values of vector of objectives are not preferable and **uncertainty is relatively low**, should have low preference;
- the aspirations to elicit global information about objective functions, and to improve the current candidate solution should not dominate each other.

The axioms which formally express these principles are presented in [6]. It is shown there that these axioms imply the single objective P-algorithm. The idea of the single-objective P-algorithm can be easily generalized to multi-objective case. Let assume that the vector Gaussian random field $\Xi(X) = (\xi_1(X), \dots, \xi_m(X))^T$, $X \in \mathbf{A}$, is accepted as a statistical model of the considered vector objective function. Assuming the components of $\Xi(X)$ independent, the improvement probability is defined by the following formula

$$\begin{aligned} & \mathbf{P}\{\Xi(X) \leq Y^{on} | \Xi(X_1) = Y_1, \dots, \Xi(X_n) = Y_n\} \\ &= \prod_{i=1}^m \mathbf{P}\{\xi_i(X) \leq y_i^{on} | \xi_i(X_1) = y_{1i}, \dots, \xi_i(X_n) = y_{ni}\} \\ &= \prod_{i=1}^m \int_{-\infty}^{y_i^{on}} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(m_i(X|\cdot) - t)^2}{2s_i^2(X|\cdot)}\right\} dt = \prod_{i=1}^m G(y_i^{on}, m_i(X|\cdot), s_i(X|\cdot)), \quad (1) \end{aligned}$$

where Y^{on} denotes the vector of objectives intended to overrun at n -th step, $m_i(X|\cdot)$ and $s_i^2(X|\cdot)$ denote the conditional mean and conditional variance of the random field $\xi_i(X)$ at the point X , and $G(\cdot, m, s)$ denotes the Gaussian (with the mean value m and the standard deviation s) cumulative distribution function. The explicit formulae of $m_i(X|X_j, y_{ij}, j = 1, \dots, n)$ and $s_i^2(X|X_j, y_{ij}, j = 1, \dots, n)$ are presented below assuming that $\xi_i(X)$ are homogeneous isotropic random fields

$$\begin{aligned} m_i(X|X_j, y_{ij}, j = 1, \dots, n) &= \mu_i + (y_{i1} - \mu_i, \dots, y_{in} - \mu_i)\Sigma^{-1}\Upsilon^T, \\ s_i^2(X|X_j, y_{ij}, j = 1, \dots, n) &= \sigma_i^2(1 - \Upsilon\Sigma^{-1}\Upsilon^T), \\ \Upsilon &= (\rho_i(X_1, X), \dots, \rho_i(X_n, X)), \quad \Sigma = \begin{pmatrix} \rho_i(X_1, X_1) & \dots & \rho_i(X_1, X_n) \\ \dots & \dots & \dots \\ \rho_i(X_n, X_1) & \dots & \rho_i(X_n, X_n) \end{pmatrix}, \quad (2) \end{aligned}$$

where μ_i , σ_i^2 , and $\rho_i(\cdot, \cdot)$ denote mean, variance and correlation function of $\xi_i(X)$. Such a generalization seems reasonable when the desirable to approach vector Y^{on} is a priori

well justified, and the closest to Y^{on} subset of the Pareto set should be approximated. It is well known that the strategy of search of the single-objective P-algorithm depends on the threshold $\Delta = \min\{y_1, \dots, y_n\} - y^{on}$: the increase of Δ increases the globality of search, and contrary, vanishing in the course of optimization Δ improves the local convergence rate. Correspondingly, by the choice of appropriate thresholds in the multi-objective P-algorithm, the size of the neighboring to Y^{on} subset of the Pareto set can be selected where the increase of the density of the computed objective vectors is desirable. For some experimental results which demonstrate the performance of the multi-objective P-algorithm we refer to [8]

Conclusions

The statistical models based single-objective GO algorithms can be generalized to the multi-objective case. Although the straightforward development of multi-objective versions is possible, the generalization of the theoretical results which substantiate single-objective algorithms is highly desirable.

Acknowledgements

The support by Agency for Science, Innovation and Technology (MITA) through the grant Nr.31V-145 is acknowledged.

References

- [1] Deb K. (2009). *Multi-Objective Optimization Using Evolutionary Algorithms*. J.Wiley, Chichester.
- [2] Miettinen K. (1999). *Nonlinear multiobjective optimization*. KAP, Boston.
- [3] Mockus J. (1988). *Bayesian approach to global optimization*. KAP, Boston.
- [4] Nakayama H., Yun Y., Yoon M. (2009). *Sequential Approximate Multiobjective Optimization Using Computational Intelligence*. Springer, Berlin.
- [5] Strongin R., Sergeyev Y. (2000). *Global Optimization with Non-convex Constraints: Sequential and Parallel Algorithms*. KAP, Boston.
- [6] Zhigljavsky A., Žilinskas A. (2008). *Stochastic Global Optimization*. Springer, NY.
- [7] Žilinskas A. (1982). Axiomatic approach to statistical models and their use in multimodal optimization theory. *Mathematical Programming*. Vol. **22**, pp. 104-116.
- [8] Žilinskas A. (2012). A statistical model-based algorithm for black-box multi-objective optimization. *International Journal of Systems Science*. Published on Internet 4 July, 2012, DOI:10.1080/00207721.2012.702244.