Lietuvos matematikos rinkinys. LMD darbai
52 tomas, 2011, 280–284

ISSN 0132-2818
www.mii.lt/LMR/

On the experimental investigation of Pareto–Lipschitzian optimization

Jonas Mockus, Justas Stašionis

Institute of Mathematics and Informatics, Vilnius University
Akademijos 4, LT-08663 Vilnius
E-mail: jmockus@gmail.com; justas.stasionis@mii.vu.lt

Abstract. A well-known example of global optimization that provides solutions within fixed error limits is optimization of functions with a known Lipschitz constant. In many real-life problems this constant is unknown.

To address that, we propose a novel method called Pareto Lipschitzian Optimization (PLO) that provides solutions within fixed error limits for functions with unknown Lipschitz constants.In the proposed approach, a set of all unknown Lipschitz constants is regarded as multiple criteria using the concept of Pareto Optimality (PO).

Keywords: Pareto, Lipschitz, Global optimization.

Introduction

The target of this paper is to discuss the results of experimental calculations. The theoretical part of PLO is described and some computing results are presented in [6]. The description of the PLO algorithm is in [6], too. We compare PLO to the existing family of the DIRECT algorithms [3, 1, 2, 4]. The DIRECT algorithms considers only a small subset of PO decisions that are selected by a heuristic rule depending on an adjustable parameter. It means that some PO decisions are preferred to others. In contrast, PLO regards all PO decisions without preferences and is naturally suited to utilize highly parallel computing.

1 Pareto–Lipschitzian Optimization (PLO)

We start to explain the optimization of Lipschitz functions with unknown constants by considering this one-dimensional example.

Suppose that the interval $D = [a, b] \in R$ is partitioned into intervals $[a_i, b_i]$, $i = 1, \dots, I$ of lengths $l_i = b_i - a_i$ with midpoints $c_i = (b_i + a_i)/2$ and the values of the function $f_\omega(x)$ are known only at the midpoints c_i. The unknown Lipschitz constants ω are regarded as different components of multiple criteria. The variables x are represented by the intervals $a_i \leqslant x \leqslant b_i$ and the function $f_\omega(x)$ is approximated by the lower bounds: $f_\omega(x) \geqslant f(c_i) - \omega l_i/2$, $a_i \leqslant x \leqslant b_i$.

Definition 1. The interval i: $a_i \leqslant x \leqslant b_i$ belonging to a compact set $D \in R$ dominates the interval j: $a_j \leqslant x \leqslant b_j$ if

$$f(c_i) - \omega\ l_i/2 \leqslant f(c_j) - \omega\ l_j/2 \quad \text{for all} \ \ \omega \in \Omega, \tag{1}$$

$$f(c_i) - \omega\ l_i/2 < f(c_j) - \omega\ l_j/2, \quad \text{for at least one } \omega \in \Omega. \tag{2}$$

Expressions (1), (2) show that the lower bound of the interval i is increasing with $f(c_i)$ and decreasing with l_i for all ω.

In [7] the set of Pareto Optimal (PO) intervals is defined as follows:

Definition 2. The interval j: $a_j \leqslant z \leqslant b_j$ belonging to a compact set $D \in R$ is called Pareto Optimal (PO), if there is no dominant interval i defined by (1), (2).

The PLO algorithm is described in [6] using corresponding theorems 1 and 2.

2 Comparison of PLO with the DIRECT algorithm

The DIRECT algorithm [3, 2] for Lipschitzian optimization with unknown constants is defined as a heuristic without any references to the theory of vector optimization or Pareto optimality. However, it can be explained in terms of PLO as well. The basic idea of DIRECT is to select (and sample within) all Potentially Optimal (PTO) intervals during an iteration. A formal definition of PTO intervals follows.

Definition 3. The interval j is said to be PTO if there exists some rate-of-change constant $\omega > 0$ such that

$$f(c_j) - \omega\ l_j/2 \leqslant f(c_i) - \omega l_i/2 \quad \text{for all } i = 1, \dots, I, \tag{3}$$

$$f(c_j) - \omega\ l_j/2 \leqslant f_{min} - \epsilon\ |f_{\min}|. \tag{4}$$

Here $\epsilon > 0$ is a constant that defines the size of the set of PTO intervals, and $f_{\min}$ is the current best value.

One can see that inequality (3) includes all the intervals which are among the best for at least one $\omega > 0$. It means that these intervals are not dominated, thus, they belong to the PO set as well. However, condition (3) defines only a subset of not dominated intervals. The further restriction of this part of PO intervals is provided by condition (4) which depends on the parameter ϵ. Thus the actual performance of DIRECT algorithm is determined by this parameter.

3 Extension to several dimensions

We define the length l_i of the interval $a_i^k \leqslant z_k \leqslant b_i^k$, $k = 1, \dots, K$ in a compact subset $D \in R^K$ as the longest length: $l_i = \max_k l_i^k$. where $l_i^k = b_i^k - a_i^k$, $k = 1, \dots, K$. The observation points c_i are in the middle $c_i^k = (b_i^k + a_i^k)/2$, $k = 1, \dots, K$. Then the definition of PO-intervals remains the same.

3.1 Sampling

The sampling procedure of the PLO algorithm [6] is similar to that of the DIRECT algorithm [2]. However, the definition of Pareto Optimal intervals, used in PLO, is different from the definition of Potentially Optimal intervals in the DIRECT algorithm.

Table 1. Comparison of three algorithms optimizing real applications.

Function	Iterations	PLO	BA	MC
EcoDuel	100	0.077	0.196	0.529
EcoDuel	1000	0.0285	0.0693	0.273
Packer	10	−2.2623	−2.2572	−2.2119
ModelTask	100	4.647	5.638	4.916

3.2 Experimental computing

Some results comparing the sequential version of PLO with other methods are in [6]. These and other calculations can be repeated independently using any of the following websites:

http://soften.ktu.lt/~mockus,
http://optimum2.mii.lt/,
http://prof.if.ktu.lt/~jonas.mockus,

section ‘Software Systems’, task ‘GMJ4: Global Optimization of many Models with PloN, nerijus version J2sdk1.6 #1’, start “Applet: gmj5.html”, select method “PloNj” and the corresponding task, click “Operations” and “Run”.

In this paper, three models that simulates real optimization problems [5] are investigated. The first one, called “Eco Duel”, in short, is about a differential game that represents the competition of two servers, using the concept of Nash equilibrium. Here eight optimization variables are used. The second five-dimensional model optimizes the “mixture” of five heuristics for packing rectangular boxes of different size into the container and is called “Packer”, in short. The third model “ModelTask” minimizes the deviation of the neuron gate model from the experimental results [8].

Table 1 illustrates the sample of PLO, BA, and MC comparison. In the all instances PLO was the best. BA was better MC with exception of the “ModelTask” model.

Figure 1 shows the results of optimal packing of hundred different rectangular packages into the rectangular container. In this example, both PLO and BA were used to optimize the mixture of five greedy heuristics, the function was the total volume of packages in the container.

Table 2 illustrates the comparison the results of Monte Carlo (MC), PLO, DIRECT obtained using standard test functions of global optimization [3, 1, 4]. MC and PLO represent methods intended for parallel computing with no adjustable parameters. The version of DIRECT is for sequential realization with one adjustable parameter ϵ selected by the method authors. This means that we compare algorithms representing different families of optimization methods.

DIRECT was better than PLO for 1 test function: Shekel, m=10. For the function Hartman-3, the results were equal. PLO was better than DIRECT for 5 functions Hartman-6, Brcos, GolPri, and SixH.

4 Summary

The theoretical novelty of PLO is the definition of the problem of Lipschitzian optimization with unknown Lipschitz constants in terms of Pareto optimality (PO).

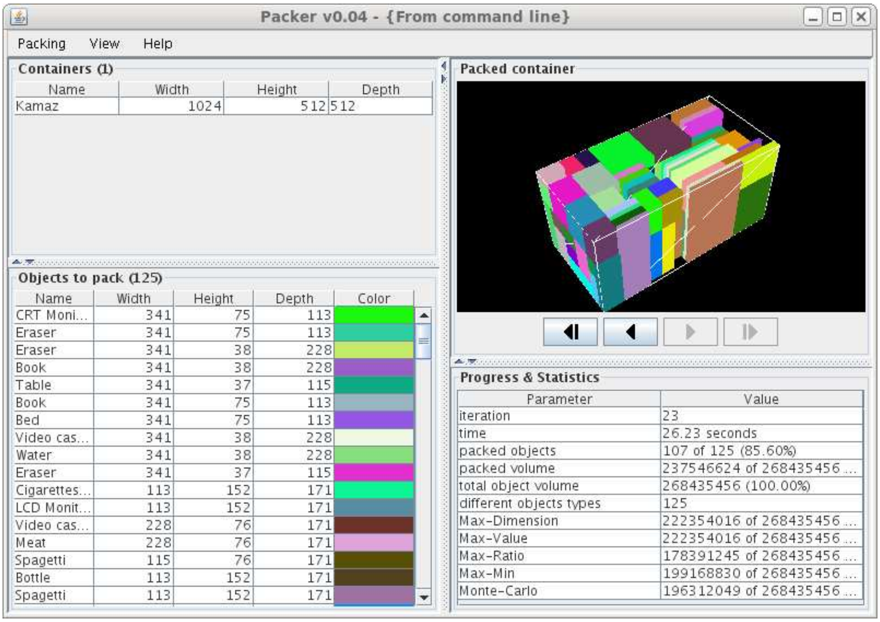

Fig. 1. Results of optimal packing using PLO.

Table 2. Comparison of three algorithms using standard test functions.

Funct	Iter	MC	DIRECT	PLO	$f(x*)$
Shekel, $m=10$	97	-1.201	-10.435	-8.54	-10.5364
Hartman-3,	83	-3.564	-3.862	-3.862	-3.86278
Hartman-6,	213	-2.505	-3.321	-3.322	-3.32237
Brcos	63	1.066	0.42	0.401	0.397887
GolPri	101	8.86	3.03	3.00	3.00
SixH	113	-0.82	-1.022	-1.027	-1.031628453
Shub2D	2883	-162.3	-184.73	-172.44	-186.7309

The computational contribution is implementation the Pareto-Lipschitzian Optimization (PLO) algorithm as the Java applet with experimental calculations illustrating the PLO efficiency for functions up to 20 variables. We expect that in the future, the extensive computer simulation with various test and real functions will reveal additional aspects of the proposed algorithm and that would be an interesting new investigation.

References

[1] M. Bjorkman and K. Holmstrom. Global optimization using the DIRECT algorithm in Matlab 1. *AMO – Adv. Mod. Opt.*, **1**:17–37, 1999.

[2] D.E. Finkel. *DIRECT optimization algorithm user guide.* Technical report, Center for Research in Scientific Computation. North Carolina State University, Raleigh, NC 27695-8205, 2003.

[3] D.R. Jones, C.D. Perttunen and B.E. Stuckman. Lipschitzian optimization without the Lipschitz constant. *J. Opt. Theor. Appl.*, **79**:157–181, 1993.

[4] G. Liuzzi, S. Lucidi and V. Piccialli. A DIRECT-based approach exploiting local minimizations for the solution of large-scale global optimization problems. *Comput. Optim. Appl.*, **45**:353–375, 2010.

[5] J. Mockus. *A Set of Examples of Global and Discrete Optimization: Application of Bayesian Heuristic Approach*. Kluwer Academic Publishers, Dordrecht–Boston–London, 2000.

[6] J. Mockus. On the Pareto-Lipschitzian optimization. *Informatica*, **22**(4), 2011.

[7] J. Mockus, W. Eddy, A. Mockus, L. Mockus and G. Reklaitis. *Bayesian Heuristic Approach to Discrete and Global Optimization*. Kluwer Academic Publishers, Dordrecht-London-Boston, ISBN 0-7923-4327-1, 1997.

[8] H. Pranevičius N. Paulauskas, M. Pranevičius and F. Bukauskas. A stochastic four-state model of contingent gating of gap junction channels containing two fast gates sensitive to transjunctional voltage. *Biophys. J.*, **96**, 2009.

REZIUMĖ

Apie eksperimentinį Pareto-Lipšico optimizacijos tyrimą

J. Mockus, J. Stašionis

Optimizavimas funkcijų su žinoma Lipšico konstatnta, užtikrinantis sprendinius duotu tikslumu, yra žinomas globaliojo optimizavimo uždavinys. Ši konstanta nėra žinoma daugelyje praktinių uždavinių. Mes tiriam naują metodą tokiems uždaviniams spręsti, pavadintą Pareto-Lipšico optimizacija (PLO). Šis metodas pateikia sprendinius duotu tikslumu funkcijoms su nežinoma Lipšico konstanta. PLO požiūriu, Lipšico konstantos nagrinėjamos kaip vektorinio kriterijaus elementai Pareto optimalumo teorijos rėmuose.

Raktiniai žodžiai: Pareto optimalumas, Lipšicas, globaliojo optimizavimo uždaviniai.