Univariate Global Optimization with Point-Dependent Lipschitz Constants

O.V. Khamisov[1,a)] and M. Posypkin[2,b)]

[1] *Melentiev Energy Systems Institute of Siberian Branch of the Russian Academy of Sciences, 664033, Irkutsk, Lermontov St., 130, Russia.*
[2] *Federal Research Center Computer Science and Control of the Russian Academy of Sciences, 119333, Moscow, Vavilova St., 40, Russia.*

[a)] Corresponding author: khamisov@isem.irk.ru
[b)] mposypkin@gmail.com

Abstract. We consider a concept of so called point-dependent Lipschitz constant. In contrast to the standard Lipschitz the value of the point-dependent Lipschitz constant can vary within the given interval. This allows us to construct a better piece-wise approximation of the objective function and as a consequence obtain a faster finding globally optimal solution.We briefly describe rules of point-dependent Lipschitz constant constructing and give an illustrative example.

Introduction

We consider univariate global optimization problem

$$\min_{x\in X} f(x), \tag{1}$$

where $X \subset \mathbb{R}$ is a closed interval $X = [\underline{x}, \overline{x}]$, function $f : [\underline{x}, \overline{x}] \to \mathbb{R}$ satisfies Lipschitz condition with a constant L

$$|f(x) - f(y)| \le L|x - y|, \;\; x, y \in X. \tag{2}$$

Deterministic univariate global optimization methods for solving problem (1) stems from seminal works of Pijavskij [1] and Shubert [2]. In these papers authors proposed to use the Lipschitzian property of a function to determine the precision of found solutions. They used simple Lipschitzian underestimations

$$\mu(x) = f(y) - L|x - y|.$$

These ideas were further developed in works of Strongin and Sergeyev [3, 4] who established an elaborated theory ("information-statistical approach") for estimating function bounds over given intervals.

Second-order Lipschitzian bounds were studied in [5, 6]. They proposed to use the following underestimation:

$$\mu(x) = f(c) + f'(c)(x - c) - L_d(x - c)^2, \tag{3}$$

where L_d is the Lipschitzian constant for the derivative. This underestimation was further improved by Sergeyev in [7]. Sergeyev introduced a smooth support function that is closer to the objective function than (3).

The further progress in a univariate global optimization was made by an important observation that a Lipschitz constant can be replaced by interval bounds on the derivatives. In [8] authors combine ideas borrowed from Pijavskij method and interval approaches. Besides new bounds the paper introduces powerful reduction rules that can significantly speed up the search process.

Another replacement of Lipschitz constant is provided by *slopes*. A slope is defined as an interval $S_f(c)$ that satisfies the following inclusion:

$$f(x) \subseteq f(c) + S_f(c) \cdot [\underline{x}, \overline{x}],$$

where c is a point within the interval $[\underline{x}, \overline{x}]$.

Clearly $S_f(c) \subseteq [\min_{x\in[\underline{x},\overline{x}]} f'(x), \max_{x\in[\underline{x},\overline{x}]} f'(x)]$. However this inclusion if often strict: slopes can provide much tighter bounds than derivative estimations. In [9, 10] efficient algorithms for evaluating slopes are proposed. Slopes are evaluated from an algebraic expression driving by rules similarly to automatic differentiation.

It worth to note powerful global optimization techniques [11, 12, 13, 14] for a multi-variate case that can serve as a source of good ideas for univariate optimization. See [15] for a good survey of such approaches.

In this paper we suggest so called point-dependent Lipschitz constant. We will say that function f satisfies point-dependent Lipschitz condition if for every $y \in X$ we have

$$|f(x) - f(y)| \le L_y|x - y| \ \forall x \in X \tag{4}$$

for some value $L_y > 0$. We also will call value L_y point-dependent Lipschitz constant. Our aim is to derive rules of constructing L_y for given Lipschitz function f and interval X in such way that $L_y \le L \ \forall y \in X$ and $L_y < L$ for points $y \in X'$, $X' \subset X$. Similar to (2) we define the point-dependent underestimator $\mu_y(x) = f(y) - L_y|x - y|$. Obviously, if $L_y < L$ then $\mu_y(x) > \mu(x)$ and efficiency of Pijavskij-type (or Shubert-type) algorithms can be improved.

It is necessary to mention similar approaches earlier used in global univariate optimization. In [16] Lipschitz constants depending on the current intervals were introduced and investigated. Another result concerning local tuning strategies is discussed in [17] and [18].

We also have to say that we do not yet apply here a convexification technique developed by C. Floudas and his collaborators (see, for example, [19]). A combination with convexification or with convex underestimation techniques would be promising and we hope to elaborate such a mixture in the nearest future.

Deriving the Point-Dependent Lipschitz Constant

For a given point $y \in X$ we want to find the smallest value L_y such that inequality (4) is correct. Let us analyze inequalities

$$f(y) - p|x - y| \le f(x) \le f(y) + p|x - y| \ \forall x \in X,$$

where p is a parameter. Consider auxiliary functions

$$\underline{\psi}(x, p, y) = f(x) - f(y) + p|x - y|, \ \overline{\psi}(x, p, y) = f(x) - f(y) - p|x - y|$$

and

$$\underline{v}(p, y) = \min_x \underline{\psi}(x, p, y), \quad \overline{v}(p, y) = \max_x \overline{\psi}(x, p, y).$$

Let $\underline{p}_y$ and $\overline{p}_y$ be the following solutions

$$\underline{p}_y \in \text{Argmin}\{p : \underline{v}(p, y) \ge 0\}, \quad \overline{p}_y \in \text{Argmin}\{p : \overline{v}(p, y) \le 0\}. \tag{5}$$

Then $L_y = \max\{\underline{p}_y, \overline{p}_y\}$. Solving problems (5) is hardly available for arbitrary function f. We suggest to solve this problem for elementary functions like $\sin, \cos, e^x$ and so on. Then, we assume, that the objective function f is constructed from a number of elementary functions by standard operations like additions, multiplications, divisions, subtractions, compositions and the like (i.e. we consider factorable functions [20]). The point-dependent Lipschitz constant of the objective function is obtained from the its elementary functions and the corresponding standard operations. Similar approach was used in [21].

Point-dependent Lipschitz constant for function sin. For function sin we have $0 \le p \le 1$. Assume that $x \le y$. In this case function v has the following form

$$\underline{v}(p, y) = -\sqrt{1 - p^2} - \sin(y) + p(y + \arccos(p))$$

(derivation of $\underline{v}$ is a standard mathematical exercise in differential calculus). Function $\underline{v}(\cdot, y)$ is increasing and strictly concave, $\underline{v}(0, y) \le 0$, $\underline{v}(1, y) \ge 0$ for any y. Hence, equation $\underline{v}(p, y) = 0$ has a unique root. The case $x \ge y$ is analysed quite similarly. So, obtaining the corresponding value $\underline{p}_y$ is reduced to solving two univariate equations. In practice, we can use some approximation of $\underline{p}_y$ to avoid finding solutions of the nonlinear equations. The problem of determining

$\overline{p}_y$ is analyzed in the same way and gives the similar result (finding solutions of two nonlinear equations). Hence, solving problems (5) for function sin is reduced to solving four nonlinear equations, solution of each equation exists and is unique.

Point-dependent Lipschitz constant for univariate convex functions. Let an interval $[\underline{x}, \overline{x}] \subset \mathbb{R}$ and a convex differentiable function $f : [\underline{x}, \overline{x}] \to \mathbb{R}$ be given. Since

$$|f(x) - f(y)| \leqslant \max_{\underline{x} \leqslant z \leqslant \overline{x}} |f'(z)||x - y|$$

and f' is a monotonously nondecreasing function the Lipschitz constant is given by

$$L_f = \max_{\underline{x} \leqslant z \leqslant \overline{x}} |f'(z)| = \max\{|f'(\underline{x})|, |f'(\overline{x})|\}.$$

Finding the point-dependent Lipschitz constant L_y different from L_f is available when $\underline{x} < y < \overline{x}$ and is based on simple geometrical considerations. Construct two secant lines

$$l_1(x) = \frac{f(y) - f(\underline{x})}{y - \underline{x}} x + \frac{y f(\underline{x}) - \underline{x} f(y)}{y - \underline{x}}, \quad l_2(x) = \frac{f(\overline{x}) - f(y)}{\overline{x} - y} x + \frac{\overline{x} f(y) - y f(\overline{x})}{\overline{x} - y}.$$

Then

$$L_y = \max \left\{ \left| \frac{f(y) - f(\underline{x})}{y - \underline{x}} \right|, \left| \frac{f(\overline{x}) - f(y)}{\overline{x} - y} \right| \right\}.$$

Note, that the suggested approach is obviously extended to finding point-dependent Lipschitz constants for concave functions.

Since many elementary functions like $\ln(x), \log(x), e^x, \frac{1}{x}, x^{2k}$ (k is a positive integer) are either convex or concave the above technique can be used for constructing point-dependent Lipschitz constants for such functions too.

Handling the Operations

Let two functions f and g be given and the corresponding point-dependent Lipschitz constant L_y^f are L_y^g are available. Then, by straightforward checking we obtain, that $L_y^{f+g} = L_y^f + L_y^g$, where L_y^{f+g} is the point-dependent Lipschitz constant for the sum $f(x) + g(x)$.

Consider now the production $f(x)g(x)$,

$$|f(x)g(x) - f(y)g(y)| = |f(x)g(x) \pm f(y)g(x) - f(y)g(y)| \leq |f(x) - f(y)| \cdot |g(x)| + |g(x) - g(y)| \cdot |f(y)| \leq$$

$$\leq \left(L_y^f \cdot M_g + L_y^g \cdot |f(y)| \right) |x - y|,$$

where $M_g = \max\{|g(x)| : x \in [\underline{x}, \overline{x}]\}$. By interchanging roles of f and g we obtain the following inequality

$$|f(x)g(x) - f(y)g(y)| \leq \left(L_y^g \cdot M_f + L_y^f \cdot |g(y)| \right) |x - y|,$$

where $M_f = \max\{|f(x)| : x \in [\underline{x}, \overline{x}]\}$. Finally, as a point-dependent Lipschitz constant L_y^{fg} for the product $f(x)g(x)$ we can take

$$L_y^{fg} = \min \left\{ L_y^f \cdot M_g + L_y^g \cdot |f(y)|, L_y^g \cdot M_f + L_y^f \cdot |g(y)| \right\}.$$

For the reciprocal $\frac{1}{f(x)}$, providing $f(x) \geq m_f > 0 \ \forall x \in [\underline{x}, \overline{x}]$, we have

$$\left| \frac{1}{f(x)} - \frac{1}{f(y)} \right| = \left| \frac{f(y) - f(x)}{f(x) f(y)} \right| \leq \frac{L_y^f}{m_g \cdot f(y)} |x - y|.$$

Hence, the corresponding point-dependent Lipschitz constant $L_y^{\frac{1}{f}} = \frac{L_y^f}{m_g \cdot f(y)}$.

A point-dependent Lipschitz constant $L_y^{f \circ g}$ for the composition $f(g(x))$ is directly derived from the definition $L_y^{f \circ g} = L_{g(y)}^f L_y^g$.

Example

In this section we give a preliminary comparison of using exact Lipschitz constant and a point-dependent Lipschitz constant within the Pijavskij method framework. The test problem was the following

$$f(x) = \sin(x) + \sin\left(\frac{10x}{3}\right) \to \min_{x\in[2.7,7.5]}.$$

The exact Lipschitz constant

$$L^f = \max_{2.7\leqslant z\leqslant 7.5} |f'(z)| = \frac{13}{3} = 4.333.$$

The tolerance $\varepsilon_f = 10^{-3}$. The described above point-dependent Lipschitz constant for function sin was used. The Pijavskij method with L^f took 132 iterations. The same methods with L^f_y took 104 iterations. The latter means that for constructing the Pijavskij lower tooth-cover the point-dependent Lipschitz constants $L^f_{x^k}$ were used at each iterations points $x^1, x^2, \ldots, x^{104}$. The average of the point-dependent Lipschitz constant $\overline{L}^f_y = 3.4668155$. The iterations improvement $\frac{132-104}{132} \cdot 100 = 21.21\%$.

A special paper will be devoted to serious computational testing on a quite a number of testing problems.

ACKNOWLEDGMENTS

We would like to thank anonymous referees for valuable comments which essentially improved the paper.
The investigation was partially supported by the RFBR grant No. 18-07-01432

REFERENCES

[1] S. Pijavskij, Optimal Decisions **2**, 13–24 (1967).
[2] B. O. Shubert, SIAM Journal on Numerical Analysis **9**, 379–388 (1972).
[3] Y. D. Sergeyev, SIAM Journal on Optimization **5**, 858–870 (1995).
[4] R. G. Strongin and Y. D. Sergeyev, *Global optimization with non-convex constraints: Sequential and parallel algorithms*, Vol. 45 (Springer Science & Business Media, 2013).
[5] L. Breiman and A. Cutler, Mathematical Programming **58**, 179–199 (1993).
[6] W. Baritompa, Journal of Global Optimization **4**, 37–45 (1994).
[7] Y. D. Sergeyev, Mathematical Programming **81**, 127–146 (1998).
[8] L. G. Casado, J. A. Martínez, I. García, and Y. D. Sergeyev, Journal of Global Optimization **25**, 345–362 (2003).
[9] D. Ratz, Journal of Global Optimization **14**, 365–393 (1999).
[10] D. Ratz, *An optimized interval slope arithmetic and its application* (Inst. für Angewandte Mathematik, 1996).
[11] Y. Evtushenko and M. Posypkin, Optimization Letters **7**, 819–829 (2013).
[12] R. Paulavičius and J. Žilinskas, *Simplicial global optimization* (Springer, 2014).
[13] V. Gergel, V. Grishagin, and R. Israfilov, Procedia Computer Science **51**, 865–874 (2015).
[14] J. D. Pintér, *Global optimization in action: continuous and Lipschitz optimization: algorithms, implementations and applications*, Vol. 6 (Springer Science & Business Media, 2013).
[15] R. Horst and P. M. Pardalos, *Handbook of global optimization*, Vol. 2 (Springer Science & Business Media, 2013).
[16] Y. D. Sergeyev, Computational Mathematics and Mathematical Physics **35**, 553–562 (1995).
[17] Y. D. Sergeyev, M. S. Mukhametzhanov, D. E. Kvasov, and D. Lera, Journal of Optimization Theory and Applications **171**, 186–208 (2016).
[18] D. E. Kvasov and Y. D. Sergeyev, Numerical Algebra, Control & Optimization **2**, 69–90 (2012).
[19] C. A. Floudas, *Deterministic global optimization. Theory, methods and applications*, Vol. 37 (Kluwer Academic Publishers, 2000).
[20] G. P. McCormick, Mathematical Programming **10**, 147–175 (1976).
[21] O. V. Khamisov, Explicit univariate global optimization with piecewise linear support functions, Proc. DOOR 2016, CEUR-WS.org, online http://ceur-ws.org/Vol-1623/papermp19.pdf.