ISSN 1392 – 124X INFORMATION TECHNOLOGY AND CONTROL, 2007, Vol. 36, No. 4

ANALYSIS OF DIFFERENT NORMS AND CORRESPONDING LIPSCHITZ CONSTANTS FOR GLOBAL OPTIMIZATION IN MULTIDIMENSIONAL CASE

Remigijus Paulavičius[1, 2], Julius Žilinskas[2, 3]

[1]*Vilnius Pedagogical University, Studentų 39, LT-2004 Vilnius*
[2]*Institute of Mathematics and Informatics, Akademijos 4, LT-08663 Vilnius*
[3]*Vilnius Gediminas Technical University, Saulėtekio 11, LT-10223 Vilnius*

Abstract. The influence of used norm and corresponding Lipschitz constant to the speed of branch and bound algorithm for multidimensional global optimization has been investigated. Lipschitz constants of different test functions for global optimization corresponding to different norms have been estimated. The test functions have been optimized using branch and bound algorithm for Lipschitz optimization with different norms. Experiments have shown that the best results are achieved when the combination of extreme (infinite and first) and Euclidean norms is used.

Keywords: Global optimization, Branch-and-bound algorithm, Lipschitz optimization, Different norms, Lipschitz constant.

1. Introduction

Global optimization is considered in this paper. Mathematically the problem is formulated as

$$f^* = \max_{x \in D} f(x),$$

where an objective function $f(x)$, $f : \mathbf{R}^n \to \mathbf{R}$, is a nonlinear function of continuous variables, $D \subseteq \mathbf{R}^n$ is a feasible region, n is the number of variables. Besides of the global optimum f^* one or all global optimizers $x^*: f(x^*)=f^*$ should be found.

Branch and bound is a technique to solve global optimization problems. Branch and bound algorithms divide the feasible region into sub-regions and detect sub-regions, which cannot contain global optimizer, by evaluating bounds for the optimum over considered sub-regions. Optimization stops when global optimizers are bracketed in small sub-regions guaranteeing the required accuracy.

Lipschitz optimization is one of the most deeply investigated subjects of global optimization. It is based on the assumption that the slope of an objective function is bounded [2]. The advantages and disadvantages of Lipschitz global optimization methods are discussed in [1], [2]. A function $f: D \to \mathbf{R}$, $D \subseteq \mathbf{R}^n$, is said to be Lipschitz if it satisfies the condition

$$|f(x) - f(y)| \le L\|x - y\|, \forall x, y \in D, \quad (1)$$

where $L > 0$ is a constant called Lipschitz constant, D is compact and $\|\cdot\|$ denotes the norm. Euclidean norm is used most often in Lipschitz optimization, but other norms can be also considered. In [5], we showed that for dimension (n = 2) combination of two extreme (infinite and first) norms gives the number of function evaluations 22% smaller than in the case when Euclidean norm is used, and for dimension (n = 3) combination of two extreme (infinite and first) and Euclidean norms gives the number of function evaluations 39% smaller than in the case when Euclidean norm is used alone.

In this paper, we investigate how the used norm and corresponding Lipschitz constant influence the speed of algorithms for global optimization in multidimensional case.

2. Branch and bound with simplicial partitions for Lipschitz optimization

A general n-dimensional simplex-based branch and bound algorithm for Lipschitz optimization has been proposed in [6]. The rules of selection, covering, branching and bounding have been justified by results of experimental investigations.

An n-dimensional simplex is the convex hull of a set of n+1 affinely independent points in the n-dimensional space. In one-dimensional space, a simplex is a segment of line, in two-dimensional space it is a triangle, in three-dimensional space it is a tetrahedron. A simplex is a polyhedron in n-dimensional space, which has the minimal number of vertices

(n+1). Therefore, if bounds on the optimum over a sub-region defined by polyhedron are estimated using function values at all vertices of the polyhedron, a simplex sub-region requires the smallest number of function evaluations to estimate bounds.

Usually, a feasible region in Lipschitz optimization is defined by a hyper-rectangle – intervals of variables. To use simplicial partitions, the feasible region should be covered by simplices. Experiments in [6] have shown that the most preferable initial covering is face to face vertex triangulation – partitioning of the feasible region into finitely many n-dimensional simplices, whose vertices are also the vertices of the feasible region.

There are several ways to divide the simplex into sub-simplices. Experiments in [6] have shown that the most preferable partitioning is subdivision of simplex into two by a hyper-plane passing through the middle point of the longest edge and the vertices whose do not belong to the longest edge.

In Lipschitz optimization the upper bound for the optimum is evaluated exploiting Lipschitz condition (1):

$$f(x) \le f(y) + L\|x-y\|.$$

It has been suggested in [6] to estimate the bounds for the optimum over the simplex using function values at one or more vertices. The lower bound for the optimum is the largest value of the function at the vertex:

$$LB(I) = \max_{x_v \in I} f(x_v),$$

where x_v is a vertex of the simplex I. The upper bound for the optimum

$$UB(I) = \min_{x_v}\left(f(x_v) + L\max_{x\in I}\|x-x_v\|\right). \quad (2)$$

In this paper the values of the function at all vertices of the simplex are used. The branch and bound algorithm may be represented by the following pseudo-code:

```
An n-dimensional hyper-rectangle D is face-to-face
vertex triangulated into set of n-dimensional simplices
I_k, k=1,...,n!
LB = -∞
UB = ∞
While (UB - LB > ε)
  Take a simplex I_j from the set of nonsolved simplices
  UB= ∞
  LB = max (LB, max_{x_v∈I_j} f(x_v))
  UB(I_j)=min(UB, min_{x_v∈I_j} ( f(x_v) + L max_{x∈I_j} ||x - x_v|| ))
  If (UB(I_j)-LB > ε)
    divide simplex I_j into two new simplices and add them to the set
  UB =UB(I_j)
```

3. Multidimensional face to face vertex triangulation of the feasible region into *n*! simplices

In our previous work [5], a 3-dimensional hyper-rectangle was face-to-face vertex triangulated into 3-dimensional simplices as it is shown in Figure 1. But such a face to face covering is not known in general case.

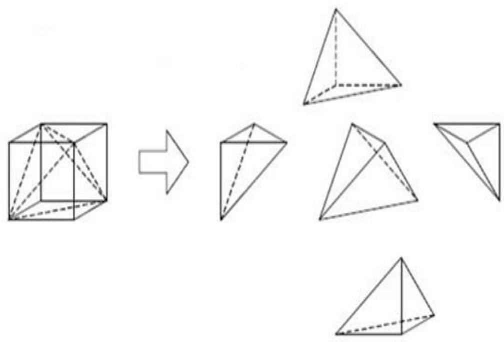

Figure 1. Old version of face to face vertex triangulation of a hyper-rectangle

Therefore we propose to use a new triangulation into n! simplices. An example of triangulation in dimension (n = 3) is shown in Figure 2. This covering is general and in such a way we can partition every dimensional hyper-rectangle.

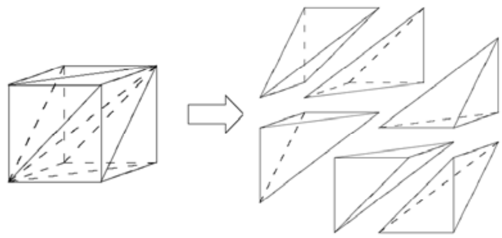

Figure 2. New version of face to face vertex triangulation of a hyper-rectangle (n = 3)

4. Norms and corresponding Lipschitz constants

Efficient algorithms for Lipschitz optimization exist for one-dimensional case. In one-dimensional case, all norms are equal. But in multidimensional case evaluated bounds depend on the norm used. It has been shown in [5] that

$$|f(x) - f(y)| \le L_p \|x-y\|_q, \quad (3)$$

where $L_p = \sup\left\{\|\nabla f(x)\|_p : x \in D\right\}$ is Lipschitz constant, $\nabla f(x) = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}\right)$ is gradient of the function $f(x)$, and $1/p+1/q=1$, $1 \le p, q \le \infty$.

The combination of two extreme and Euclidean norms is denoted by:

$$L_{1,\infty,2}\|x-x_v\|_{\infty,1,2}\left(UB_{1,\infty,2}\right)= \\ =\min\left\{L_1\|x-x_v\|_\infty,\ L_\infty\|x-x_v\|_1,\ L_2\|x-x_v\|_2\right\}. \quad (4)$$

Various test functions for global optimization from [1] and [4] have been used in our experiments for (n = 3) and from [3] and [4] for (n > 3) dimensional functions. Lipschitz constants have been estimated using Theorem 1 in [5]. The extreme (first and infinite) and Euclidean Lipschitz constants for (n > 3) dimensional functions are shown in Table 1. For (n > 3) function names and numbers from [3] and [4] are used. Test functions with (n = 3) are numbered according to [1], [4].

Table 1. Test functions and estimated Lipschitz constants for multivariate Lipschitz optimization

Function number	Lipschitz function	Domain	L_1	L_2	L_∞
1.	Shekel 5 [3]	$[0,10]^4$	204.1	102.4	56.1
2.	Shekel 7 [3]	$[0,10]^4$	300.1	151.5	86.1
3.	Shekel 10 [3]	$[0,10]^4$	408.2	204.5	110.8
4.	Levy No. 9 [3]	$[-10,10]^4$	26.1	14.4	8.3
5.	Levy No. 15 [3]	$[-10,10]^4$	1273.2	1196.4	1195.5
6.	Schwefel No. 1.2 [3]	$[-5,10]^4$	311.3	170.1	120.1
7.	Powell [3]	$[-4,5]^4$	11338.3	7245.8	5150.1
8.	Generalized Rosenbrock [4]	$[-4,4]^4$	72224.5	60401.7	36026.0
9.	Levy No. 16 [3]	$[-5,5]^5$	422.9	370.7	369.7
10.	Generalized Rosenbrock [4]	$[-5,5]^5$	264567	130914	66214
11.	Levy No. 10 [3]	$[-10,10]^5$	34.4	16.6	8.25
12.	Schwefel No. 3.7 [3]	$[-0.5,0.4]^5$	0.976	0.044	0.019
13.	Levy No. 17 [3]	$[-5,5]^6$	421.7	358.6	357.4
14.	Generalized Rosenbrock [4]	$[-6,6]^6$	548484	241742	109334
15.	Levy No. 18 [3]	$[-5,5]^7$	433.1	358.9	357.5
16.	Levy No. 12 [3]	$[-10,10]^{10}$	75.6	24.8	8.25

Table 2. Numbers of function evaluations for n=3

		Old covering	New covering			
Test function	ε	$L_{1,\infty,2}\|x-y\|_{\infty,1,2}$	$L_{1,\infty,2}\|x-y\|_{\infty,1,2}$	$L_\infty\|x-y\|_1$	$L_2\|x-y\|_2$	$L_1\|x-y\|_\infty$
20. [1]	10.6	75163	49411	49411	71800	84938
21. [1]	0.369	8866	2248	24143	7488	2248
23. [1]	41.65	96673	22065	274625	145381	22065
24. [1]	3.36	61492	28451	33848	29347	56613
25. [1]	0.0506	20335	9629	23361	10065	19161
26. [1]	4.51	18679	12023	12271	24137	28265
5. [4].	5000.0	103694	58163	72323	60472	64464

5. Results of experiments

5.1. Comparison of three -dimensional test functions with different partitions

The speed of global optimization has been estimated using the number of function evaluations. For three-dimensional functions, the influence of covering (either new or old) has been investigated. The results are shown in Table 2. The number of function evaluations with new covering is smaller from 34% to 75%.

5.2. Optimization of multidimensional test functions

Let us compare bounds evaluated using combination of norms (4) with bounds evaluated using single different norm. Let us start with n=4. Let us assume that the feasible region is unit cube D=[0,1]×[0,1]×[0,1] ×[0,1], test function is *Shekel 5* and simplex I_{ABCDE}: A(0,0,0,0), B(0,0,0,1), C(0,0,1,1), D(0,1,1,1), E(1,1,1,1).

It has been shown in [5] that $L_p = \sup\left\{\left\|\nabla f(x)\right\|_p : x \in D\right\}$, therefore, points of the feasible region, where the norm of gradient of the objective function is the largest, have been found, see Table 3. For *Shekel 5* function:

$$\left|f'_{x_1}\right| = 56.1, \left|f'_{x_2}\right| = 52, \left|f'_{x_3}\right| = 52, \left|f'_{x_4}\right| = 44 .$$

Therefore:

$$L_1 = 56.1 + 52 + 52 + 44 = 204.1 ,$$
$$L_\infty = \mathbf{max}\{56.1, 52, 52, 44\} = 56.1 ,$$
$$L_2 = \left(56.1^2 + 52^2 + 52^2 + 44^2\right)^{1/2} = 102.4 .$$

Table 3. The largest values of derivatives for n=4 and n=5

Function number	Lipschitz function	$\left\|f'_{x_1}\right\|$	$\left\|f'_{x_2}\right\|$	$\left\|f'_{x_3}\right\|$	$\left\|f'_{x_4}\right\|$	$\left\|f'_{x_5}\right\|$
1.	Shekel 5 [3]	56.1	52.0	52.0	44.0	–
2.	Shekel 7 [3]	58.0	80.0	86.14	76.0	–
3.	Shekel 10 [3]	100.0	93.2	104.2	110.8	–
4.	Levy No. 9 [3]	8.25	8.25	8.25	1.375	–
5.	Levy No. 15 [3]	1195.5	22.0	22.0	33.7	–
6.	Schwefel No. 1.2 [3]	120.08	100.6	60.4	30.2	–
7.	Powell [3]	5150.1	4970.0	1126.2	92.0	–
8.	Generalized Rosenbrock [4]	36026.0	36010.0	32214.5	4000.0	–
9.	Levy No. 16 [3]	369.68	17.25	12.0	12.0	12.0
10.	Generalized Rosenbrock [4]	66214	66012	66012	60328	6000
11.	Levy No. 10 [3]	8.25	8.25	8.25	8.25	1.375
12.	Schwefel No. 3.7 [3]	0.0195	0.0195	0.0195	0.0195	0.0195

Now we can evaluate different *UB* in all vertices of simplex I_{ABCDE} :

$$UB_{1\infty}(A) = f(A) + L_1\|A - E\|_\infty = f(A) + 204.1 \cdot 1,$$
$$UB_{1\infty}(B) = f(B) + L_1\|B - E\|_\infty = f(B) + 204.1 \cdot 1,$$
$$UB_{1\infty}(C) = f(C) + L_1\|C - E\|_\infty = f(C) + 204.1 \cdot 1,$$
$$UB_{1\infty}(D) = f(D) + L_1\|D - A\|_\infty = f(D) + 204.1 \cdot 1,$$
$$UB_{1\infty}(E) = f(E) + L_1\|E - A\|_\infty = f(E) + 204.1 \cdot 1,$$
$$UB_{22}(A) = f(A) + L_2\|A - E\|_2 = f(A) + 102.4 \cdot 2,$$
$$UB_{22}(B) = f(B) + L_2\|B - E\|_2 = f(B) + 102.4 \cdot \sqrt{3},$$
$$UB_{22}(C) = f(C) + L_2\|C - E\|_2 = f(C) + 102.4 \cdot \sqrt{2},$$
$$UB_{22}(D) = f(D) + L_2\|D - A\|_2 = f(D) + 102.4 \cdot \sqrt{3},$$
$$UB_{22}(E) = f(E) + L_2\|E - A\|_2 = f(E) + 102.4 \cdot 2,$$
$$UB_{\infty 1}(A) = f(A) + L_\infty\|A - E\|_1 = f(A) + 56.1 \cdot 4,$$
$$UB_{\infty 1}(B) = f(B) + L_\infty\|B - E\|_1 = f(B) + 56.1 \cdot 3,$$
$$UB_{\infty 1}(C) = f(C) + L_\infty\|C - E\|_1 = f(C) + 56.1 \cdot 2,$$
$$UB_{\infty 1}(D) = f(D) + L_\infty\|D - A\|_1 = f(D) + 56.1 \cdot 3,$$
$$UB_{\infty 1}(E) = f(E) + L_\infty\|E - A\|_1 = f(E) + 56.1 \cdot 4.$$

Suppose function values at all vertices are similar:

$$f(A) \approx \ldots \approx f(E) = f(x_v) .$$

Because

$$UB(I) = \min_{x_v}\left(f(x_v) + L \max_{x \in I}\|x - x_v\|\right)$$

we get:

$$UB_{1\infty}(I_{ABCDE}) = f(x_v) + 204.1 \cdot 1 = f(x_v) + 204.1,$$
$$UB_{22}(I_{ABCDE}) = f(x_v) + 102.4 \cdot \sqrt{2} = f(x_v) + 144.82,$$
$$UB_{\infty 1}(I_{ABCDE}) = f(x_v) + 56.1 \cdot 2 = f(x_v) + 112.2 .$$

Therefore the best results for this function are got when $UB_{\infty 1}$ is used. But it depends on the values of function derivatives. If all derivatives are aproximately equal (*Schwefel No. 3.7*, *Shekel 5, Shekel 7, Shekel 10*), then the best *UB* is got, when $UB_{\infty 1}$ is used. If the value of one derivative is much larger than others (*Levy No. 15*, *Levy No. 16*), then the best *UB* is got when $UB_{1\infty}$ is used. In others cases, $UB_{\infty 1}$ or UB_{22} got the best upper bound.

The results of experiments for n=4,…,10 are shown in Table 4. The results for n=4 functions, with better precision are shown in Table 5.

None of the single norm and corresponding Lipschitz constant is the best for all test functions. The best results have been achieved using the combination of two extreme norms and Euclidean norm.

6. Conclusions

In this paper a general algorithm for multidimensional Lipschitz global optimization is tested. Test functions of different dimensionality (n = 4, 5, …, 10) have been used for experimental investigation of the algorithm.

The combination of two extreme (infinite and first) and Euclidean norms gives the best results for Lipschitz optimization. When $n = 4$, the number of function evaluations on the average is by 13% smaller when the combination is used than when Euclidean norm is used. When $n = 5$, the number of function evaluations on the average is by 37% smaller that when Euclidean norm is used. For n –dimensional functions, where $n = 6, \ldots, 10$, functions we got about 60% smaller number of function evaluations.

Table 4. Number of function evaluations for $n > 3$

Test function	Precision	$L_{1,\infty,2}\|x-y\|_{\infty,1,2}$	$L_{1,\infty}\|x-y\|_{\infty,1}$	$L_\infty\|x-y\|_1$	$L_2\|x-y\|_2$	$L_1\|x-y\|_\infty$
1.	$2L_2$	29212	29212	29212	30956	227107
2.	$2L_2$	29212	29212	29212	30956	227107
3.	$2L_2$	29212	29212	29212	30956	227107
4.	$2L_2$	30515	34650	35346	40530	180168
5.	$2L_2$	227107	227107	>4000000	440381	227107
6.	$2L_2$	50161	94752	99737	51049	202409
7.	$2L_2$	2786	5717	7040	3016	9930
8.	$2L_2$	1917	9689	138108	1958	36816
9.	$2L_2$	296147	296998	>4000000	877589	296998
10.	$2L_2$	271676	281459	281547	537788	>2500000
11.	$2L_2$	606470	679873	680855	885330	>2500000
12.	$2L_2$	33	33	33	33	33
13.	$3L_2$	208904	208904	>1200000	>1200000	208904
14.	$3L_2$	479673	498411	499683	>1200000	>1200000
15.	$5L_2$	78124	78124	>1200000	305034	78124
16.	$8 L_2$	79744	89314	89443	116411	>1200000

Table 5. Number of function evaluations for n=4 with precision L_2

Test function	$L_{1,\infty,2}\|x-y\|_{\infty,1,2}$	$L_{1,\infty}\|x-y\|_{\infty,1}$	$L_\infty\|x-y\|_1$	$L_2\|x-y\|_2$	$L_1\|x-y\|_\infty$
1.	293786	293786	293786	527231	3839611
2.	537243	537243	537243	563667	3839611
3.	293848	293848	293848	563667	3839611
4.	193057	225503	231271	264649	1252493
5.	3432907	3432907	>4000000	>4000000	3432907
6.	509388	934936	995994	517490	2001558
7.	39393	87288	104410	40349	151266
8.	95391	135662	138108	116939	468363

Acknowledgements

The research is supported by Lithuanian State Science and Studies Foundation within the project B-03/2007 "Global optimization of complex systems using high performance computing and GRID technologies" and the NATO Reintegration grant CBP.EAP.RIG.981300.

References

[1] **P. Hansen, B. Jaumard.** Lipshitz optimization. *In: R. Horst, P. Pardalos (eds.). Handbook of Global Optimization. Kluwer Academic Publishers*, 1995, 404-493.

[2] **R. Horst, P.M. Pardalos, N.V. Thoai.** Introduction to Global Optimization. *Kluwer Academic Publishers*, 1995.

[3] **C. Jansson, O. Knüppel.** A Global Minimization Method: The Multi-Dimensional Case. *Technical Report, TU Hamburg-Harburg*, 1992.

[4] **K. Madsen, J. Žilinskas.** Testing Branch-and-Bound Methods for Global Optimization. *Technical Report* IMM-REP-2000-05, *Technical University of Denmark*, 2000.

[5] **R. Paulavičius, J. Žilinskas.** Analysis of different norms and corresponding Lipschitz constants for global optimization. *Technological and Economic Development of Economy*, 12 (4), 2006, 301-306.

[6] **J. Žilinskas.** Optimization of Lipschitzian functions by simplex-based branch and bound. *Information Technology and Control*, *Vol.*14, *No.*1, 2000, 45-50.

Received August 2007.