CREAT. MATH. INFORM. **24** (2015), No. 2, 113 - 123
Online version at http://creative-mathematics.ubm.ro/
Print Edition: ISSN 1584 - 286X Online Edition: ISSN 1843 - 441X

A numerical study on the robustness and efficiency of the PL homotopy algorithm for solving unconstrained optimization problems

ANDREI BOZANTAN [1] and VASILE BERINDE [1,2]

ABSTRACT. Our aim in this paper is to illustrate the relevance of the fixed point piecewise-linear homotopy algorithm for solving unconstrained optimization problems. The numerical tests are performed by using an implementation of the piecewise-linear homotopy algorithm in the modern programming language C#, as described previously in [Bozantan, A., *An implementation of the piecewise-linear homotopy algorithm for the computation of fixed points*, Creat. Math. Inform., **19** (2010), No. 2, 140–148] and [Bozantan, A. and Berinde, V., *Applications of the PL homotopy algorithm for the computation of fixed points to unconstrained optimization problems*, Creat. Math. Inform., **22** (2013), No. 1, 41–46]. As shown by the numerical experiments done on a set of classic test functions in optimization theory, the PL homotopy algorithm appears to be more reliable than the classical Newton's method and some other important methods for finding local or global minima.

1. INTRODUCTION

We consider the unconstrained optimization problem

$$\min f(x), x \in \mathbb{R}^n, \tag{1.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable. There exists many (Newton type) iterative methods in literature for solving problem (1.1), see for example [12], [16], [18]. Such an iterative method produces a monotone or non monotone sequence $x_0, x_1, x_2, ...$, where x_{k+1} is generated from x_k, the current direction d_k, and the stepsize α_k by the rule

$$x_{k+1} = x_k + \alpha_k d_k. \tag{1.2}$$

Amongst the most reliable and largely used iterative methods in unconstrained optimization, Newton's method and Newton type methods play an important role, due to the fact that they allow us to identify by a certain procedure the search directions.

In order to decide which one of these methods should be more appropriate for a certain problem, it would be desirable to know *a priori*, if possible, a scale of the most efficient and robust techniques, for as many as possible classes of objective functions.

Efficiency is an important feature of any iterative procedure, since in concrete problems for more than three or four variables trial and error becomes impractical because, in some regions, the optimization algorithm may progress very slowly toward the optimum, requiring excessive computer time.

Robustness, i.e., the ability to achieve a solution, is equally or even more important because a general nonlinear function is unpredictable in its behavior: there may be local maxima or minima, saddle points, regions of convexity, concavity, and so on. Therefore,

Received: 22.03.2015. In revised form: 30.06.2015. Accepted: 22.09.2015
2010 *Mathematics Subject Classification.* 47H10, 47J25, 49M15, 49M30.
Key words and phrases. *Piecewise-linear homotopy algorithm, unconstrained optimization problem, Newton's method, test function.*
Corresponding author: V. Berinde; vberinde@ubm.ro

it is of great theoretical and practical importance to draw an extensive experience in testing optimization algorithms for unconstrained functions to evaluate their efficiency and robustness.

The main aim of this paper is to illustrate the relevance of the piecewise-linear homotopy algorithm (which has been designed and implemented for fixed point approximation, see [4], [6] and [5]), for solving unconstrained optimization problems and to compare its performances to well known iterative methods: Newton's method, Broyden-Fletcher-Goldfarb-Shanno, conjugate gradient method, and nonlinear conjugate gradient method. This empirical study is done on a set of representative test functions taken from literature, see [13], [14], [22], [23], [24].

2. PRELIMINARIES

This new approach to unconstrained optimization problems is based on the fact that algorithm (1.2) can be regarded as a particular case of a classic fixed point iterative method, that is, of the Picard iteration or successive approximations method associated to a certain nonlinear fixed point equation

$$x = Tx, \tag{2.3}$$

where T is a given self operator of a space X. Suppose X and T are such that the equation (2.3) has at least one solution (usually called a *fixed point* of T). A typical situation of this kind is illustrated by the well known Brouwer's fixed point theorem, see [17].

Theorem 2.1. *Every continuous mapping f from a convex compact subset K of a Euclidean space to K itself has a fixed point.*

Under the assumptions of Theorem 2.1, the Picard iteration associated to (2.3), defined by $x_0 \in X$ and

$$x_{n+1} = Tx_n,\ n = 0, 1, 2, \ldots, \tag{2.4}$$

does not converge, in general, even though in many cases (e.g., for contractive type mappings, see [2]) it is a useful method to solve nonlinear fixed point equations.

Any contraction mapping is continuous but the reverse is not true. This is the reason why several authors tried to find specific algorithms that could be successfully used to compute fixed points of continuous but not contractive mappings.

In this context, in 1967 Herbert Scarf proposed a method for approximating fixed points of continuous mappings [25]. The algorithm proposed by Scarf, which is also a numerically implementable constructive proof of the Brouwer fixed point theorem, has its origins in the Lemke-Howson complementary pivoting algorithm for solving linear complementarity problems [20]. Beside the generalization and applications in fixed point theory, the Lemke-Howson algorithm is also famous for its applications in finding Nash equilibrium points for bimatrix games.

Several improvements to the algorithm developed by Scarf were made by Terje Hansen in 1967, see [26] and by Harold W. Kuhn in 1968 [19]. But the decisive advancements came in 1972, when Eaves [10] and then Eaves and Saigal [11] described a piecewise-linear (PL) homotopy deformation algorithm as an improvement for the algorithm proposed by Scarf. Another PL algorithm, related to the one proposed by Eaves and Saigal, was presented by Orin H. Merrill in 1972 [21]. The main practical advantage of the PL homotopy methods is that they don't require smoothness of the underlying map, and in fact they can be used to calculate fixed points of set-valued maps. Although PL methods can be viewed in the more general context of complementary pivoting algorithms, they are usually considered in the special class of homotopy or continuation methods [1].

Since the piecewise-linear homotopy method could be in particular applied to solve optimization problems, the main aim of this paper is to test numerically the efficiency and robustness of the PL homotopy algorithm in the case of unconstrained optimization problems and to compare it with some of the well known and widely used methods in optimization.

The tests reported in Section 4 of this paper clearly illustrate the fact that, for almost all test functions we have considered, but especially for the non differentiable ones, the piecewise-linear homotopy method is more robust (even though not always more efficient) than Newton's method, Broyden-Fletcher-Goldfarb-Shanno, conjugate gradient method, nonlinear conjugate gradient method. This reveals the usefulness and numerical potential of the PL homotopy algorithm for solving unconstrained optimization problems.

3. Piecewise-linear homotopy algorithms

A detailed description of the implementation and use of the piecewise-linear homotopy algorithm for solving some unconstrained optimization problems is given in [4] and [5]. To make this paper self contained we present some excerpts in the following, for the convenience of the reader.

The homotopy methods are useful alternatives and aides for the Newton methods in solving systems of n nonlinear equations in n variables:

$$F(x) = 0, \quad F : \mathbb{R}^n \to \mathbb{R}^n. \tag{3.5}$$

mainly when very little a priori knowledge regarding the zero points of F is available and so, a poor starting value could cause a divergent Newton iteration sequence. The idea of the homotopy method is to consider a new function $G : \mathbb{R}^n \to \mathbb{R}^n$, related to F, with a known solution, and then to gradually deform this new function into the original function F. Typically one can define the convex homotopy:

$$H(x,t) = t \cdot G(x) + (1-t) \cdot F(x) \tag{3.6}$$

and can try to trace the implicitly defined curve

$$H^{-1}(0) = \{x \in \mathbb{R}^n \mid \exists\, t \in [0,1] \text{ such that } H(x,t) = 0\} \tag{3.7}$$

from a starting point $(x_0, 1)$ to a solution point $(x^*, 0)$. The implicit function theorem ensures that the set $H^{-1}(0)$ is at least locally a curve under the assumption that $(x_0, 1)$ is a regular value of H, i.e. the Jacobian $H'(x_0, 1)$ has full rank n. However, because there is no smoothness condition on F, a more complex approach involving piecewise-linear approximations is needed.

We define the "refining" triangulation J_3 of $\mathbb{R}^n \times (0,1]$ such that the vertices of this triangulation are given by the set of points:

$$J_3^0 = \{(v_1, \ldots, v_{n+1}) \mid v_{n+1} = 2^{-k}, k \in \mathbf{N} \text{ and } \frac{v_i}{v_{n+1}} \in \mathbb{Z}\}.$$

So, every $(n+1)$-simplex of this triangulation is contained in some slab $\mathbb{R}^n \times [2^{-k}, 2^{-k-1}]$, $k \in \mathbf{N}$.

Let $\sigma = [v_1, v_2, \ldots, v_{n+1}, v_{n+2}] \in J_3$ be an $(n+1)$-simplex and let $\pi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ be the following canonical projection: $\pi(x,t) = t$. We define the level of σ as $\max_{i=\overline{1,n+2}} \pi(v_i)$, which is the maximum of the last co-ordinates of all vertices of σ. We call J_3 a refining triangulation of $\mathbb{R}^n \times \mathbb{R}$ because the diameter of σ tends to zero as the level of σ tends to zero.

We define the piecewise linear homotopy map H_{J_3} which interpolates H on the vertices of the given refining triangulation J_3:

- $H_{J_3}(x,1) = G(x)$, $H_{J_3}(x,0) = F(x)$;
- $H_{J_3}(x,t) = \sum_{i=1}^{n+2} \lambda_i H(v_i,t)$, where: (v_i,t) are vertices of $\sigma \in J_3$ and $(x,t) = \sum_{i=1}^{n+2} \lambda_i (v_i,t), \quad \sum_{i=1}^{n+2} \lambda_i = 1, \lambda_i \geq 0.$

The algorithm will trace the unique component of the polygonal path $H_{J_3}^{-1}(0)$ which contains $(x_0,1)$, with nodes on the n-faces of the triangulation J_3.

The algorithm starts with the unique simplex σ_0 which contains the initial point $(x_0,1)$. Then, for each $i = 0,1,\ldots$ it will perform the following steps in a loop:

- It will trace the restriction of $H_{J_3}^{-1}(0)$ to the current simplex σ_i, from the point (x_i,t_i) and finds the intersection point (x_{i+1},t_{i+1}) with some other facet of σ_i. This step is called "door-in-door-out step", see [1]. Sometimes this step is also called linear programming step because it involves the solving of linear equations in a manner typical for linear programming methods.
- It performs a pivoting step, which means to find the new simplex σ_{i+1} which is adjacent to the current simplex and which contains the point (x_{i+1},t_{i+1}). This step is usually performed using only a few operations which define the pivoting rules of the triangulation.

The generated sequence $(x_0,1),(x_1,t_1),\ldots$ will converge to a solution $(x^*,0)$ of the homotopy map H, that is, to a solution x^* of the equation $F(x) = 0$.

4. Numerical tests on the robustness and efficiency of PL homotopy algorithm for unconstrained optimization problems

In order to find local or global minima of a nonlinear multivariable function $F : \mathbb{R}^n \to \mathbb{R}$, the following homotopy map is defined and used in the PL homotopy algorithm [11]:

$$H : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n,\ H(x,t) = \begin{cases} x - x_0 & \text{for } t \leq 0, \\ \nabla F(x) & \text{for } t > 0. \end{cases} \tag{4.8}$$

The numerical experiments were performed for the case $n = 2$ (Example 4.1, Example 4.2, Example 4.3) and for $n = 10$ in Example 4.4 (but the graph is done for $n = 2$, too). All the numerical results are computed to a maximal final error of $1e-4$.

We shall use the following symbols in the tables that synthesise the numerical experiments we have done:

- BFGS = Broyden-Fletcher-Goldfarb-Shanno method (a quasi-Newton method)
- CG = conjugate gradient method
- NCG = nonlinear conjugate gradient method
- PLH = piecewise linear homotopy algorithm
- ∞ = method does not converge
- $\neq$ = method converges, but not to the correct global/local minimum

Example 4.1. (**The sixth Bukin function**)

$$F(x,y) = 100 \cdot \sqrt{|y - 0.01 \cdot x^2|} + 0.01 \cdot |x + 10|,\ (x,y) \in \mathbb{R}^2.$$

The global minimum value of F is $F(-10,1) = 0$. It is usually evaluated on the rectangle $(x,y) \in [-15,-5] \times [-3,3]$. The sixth Bukin function has many local minima, all of which lie in a ridge. Note also that F is not differentiable on the set $\{(-10,b) : b \in \mathbb{R}\} \cup \{(10\sqrt{|b|},b) : b \in \mathbb{R}\}$.

From the table above we note that all algorithms taken into consideration in our comparison study for the PL homotopy algorithm do not converge in the case of the starting points $(-15,0),(-10,0),(-11,3),(-11,-3)$ (and, of course, for many other starting points). This is mainly due to the fact that the objective function F is not differentiable at some points.

The merit of the PL homotopy algorithm is that it converges in all those cases and reaches the minimal value.

Example 4.2. (**Easom function**)

$$F(x,y) = -\cos(x)\cdot\cos(y)\cdot e^{-(x-\pi)^2-(y-\pi)^2},\ (x,y) \in \mathbb{R}^2.$$

The global minimum value of F is $F(\pi,\pi) = -1$ within $-100 \leq x,y \leq 100$. It has many local minima.

Example 4.3. (**Rosenbrock function, also known as banana function**)

$$F(x,y) = (1-x)^2 + 100(y-x^2)^2,\ (x,y) \in \mathbb{R}^2.$$

The function is unimodal, and the global minimum value, $F(1,1) = 0$, lies in a narrow, parabolic valley. Usually it is evaluated within the rectangle $-5 \leq x,y \leq 5$.

Example 4.4. (**Schwefel function**)

$$F(x_1,\ldots,x_n) = \sum_{i=1}^{n} -x_i \cdot \sin\left(\sqrt{|x_i|}\right)$$

We will test the optimization algorithms for $n = 10$, in the hypercube $x_i \in [-10,10]\,, i = 1..n$. The function has a single minimum in this input domain which is:

$$F\left(\frac{5\pi}{3},\ldots,\frac{5\pi}{3}\right) = \frac{-7.89\cdot n}{2}.$$

Note also that the Schwefel function F is not differentiable on the set $\{(0,a_2,\ldots,a_n) : a_2,\ldots,a_n \in \mathbb{R}\} \cup \{(a_1,0,a_3,\ldots,a_n) : a_1,a_3,\cdots \in \mathbb{R}\} \cup \ldots$.

We shall use the following starting points for comparing the algorithms:

$$c_1 = (5,5,5,5,5,5,5,5,5,5), c_2 = (5,5,5,5,5,5,5,5,5,1), c_3 = (2,2,2,2,2,2,2,2,2,2),$$
$$c_4 = (2,2,2,2,2,2,2,2,2,1), c_5 = (1,1,1,1,1,1,1,1,1,1), c_6 = (1,1,1,1,1,1,1,1,1,0),$$
$$c_7 = (0,0,0,0,0,0,0,0,0,0).$$

5. Conclusions and future work

We tested numerically the efficiency and robustness of the PL homotopy algorithm for some typical unconstrained optimization problems and compared it with some of the well known and widely used methods in optimization: Newton's method, Broyden-Fletcher-Goldfarb-Shanno, conjugate gradient method, and nonlinear conjugate gradient method.

The numerical results clearly illustrate the fact that, for almost all test functions we have considered, but especially for the non differentiable ones, the piecewise-linear homotopy method is more robust (even though not always more efficient) than Newton's

method, Broyden-Fletcher-Goldfarb-Shanno, conjugate gradient method, and nonlinear conjugate gradient method. The main generic advantages of the PL homotopy methods is that they don't require smoothness of the underlying map. This fact is clearly illustrated by the numerical results in Examples 4.1 and 4.4. Also another important feature of these methods is that they can be applied when no a priori knowledge regarding the solutions of the system to be solved is available, and thus to choose a suitable starting point for the iterative method.

This study reveals the usefulness as well as the numerical potential of the PL homotopy algorithm for solving unconstrained optimization problems and is thus encouraging us to continue the study by considering other important optimization test functions from literature ([13], [14], [22], [23], [24].

Acknowledgement. This work was partially financed by European Social Fund through the Sectoral Operational Programme Human Resources Development 2007-2013. The second author's research was supported by the Grants PN-II-RU-TE-2011-3-239 and PN-II-ID-PCE-2011-3-0087 of the Romanian Ministry of Education and Research.

REFERENCES

[1] Allgower, E. L. and Georg, K., *Introduction to numerical continuation methods*, Springer-Verlag, Berlin, 1990
[2] Berinde, V., *Iterative Approximation of Fixed Points*, Springer, Berlin Heidelberg New York, 2007
[3] Bongartz, I., Conn, A. R., Gould, N. I. M. and Toint, P. L., *CUTE: Constrained and unconstrained testing environments*, ACM Trans. Math. Software, **21** (1995), 123–160
[4] Bozantan, A., *An implementation of the piecewise-linear homotopy algorithm for the computation of fixed points*, Creat. Math. Inform., **19** (2010), No. 2, 140–148
[5] Bozantan, A., *New implementations of fixed point iterative algorithms and applications to nonsmooth optimization*, PhD Thesis, North University Center at Baia Mare, Technical University of Cluj-Napoca, 2014
[6] Bozantan, A. and Berinde, V., *Applications of the PL homotopy algorithm for the computation of fixed points to unconstrained optimization problems*, Creat. Math. Inform., **22** (2013), No. 1, 41–46
[7] Bozantan, A. and Berinde, V., *About the implementation and some applications of the FIXPOINT software mini-package*, Creat. Math. Inform., **23** (2014), No. 1, 41–50
[8] Cegielski, A., *Iterative Methods for Fixed Point Problems in Hilbert Spaces*, Lecture Notes in Mathematics, **2057**, Springer-Verlag, Berlin, 2013
[9] Chidume C.E. and Măruşter, Şt., *Iterative methods for the computation of fixed points of demicontractive mappings*, J. Comput. Appl. Math., **234** (2010), 881–882
[10] Eaves, C. B., *Homotopies for computation of fixed points*, Mathematical Programming, **3** (1972), No. 1, 1–22
[11] Eaves, C. B. and Saigal, R., *Homotopies for computation of fixed points on unbounded regions*, Mathematical Programming, **3** (1972), No. 1, 225–237
[12] Edgar, T. F., Himmelblau, D. M., Lasdon, L. S., *Optimization of Chemical Processes*, Second Edition, McGraw-Hill, 2001
[13] Floudas, C. A., Pardalos, P. M., Adjiman, C. S., Esposito, W. R., Gumus, Z. H., Harding, S. T., Klepeis, J. L., Meyer, C. A. and Scheiger, C. A., *Handbook of Test Problems in Local and Global Optimization*, Springer, 1999
[14] Hedar, A., *Test function web pages*, http://www-optima.amp.i.kyoto-u.ac.jp/member/student/hedar/Hedar files/TestGO files/Page364.htm
[15] Herings, P. J.-J. and Peeters, R., *Homotopy methods to compute equilibria in game theory*, Economic Theory, **42** (2010), No. 1, 119–156
[16] Güler, O., *Foundations of Optimization*, Graduate Texts in Mathematics, **258**, Springer, 2010
[17] Istratescu, V. I., *Fixed Point Theory. An Introduction*, Kluwer Academic Publishers, 2001
[18] Kelley, C. T., *Iterative Methods for Optimization*, SIAM, Philadelphia, 1999
[19] Kuhn, H. W., *Simplicial approximation of fixed points*, Proceedings of the National Academy of Sciences of the United States of America, **61** (1968), No. 4, 1238
[20] Lemke, C. E., *Bimatrix equilibrium points and mathematical programming*, Management science, **11** (1965), No. 7, 681–689
[21] Merrill, O. H., *Applications and extensions of an algorithm that computes fixed points of certain upper semi-continuous point to set mappings*, Ph.D. thesis, University of Michigan., Ann Arbor, Michigan, 48106, USA (1972)

[22] Molga, M. and Smutnicki, C., *Test functions for optimization needs*, http://www.zsd.ict.pwr.wroc.pl/files/docs/functions.pdf
[23] Moré, J. J., Garbow, B. S. and Hillstrom, K. E., *Testing unconstrained optimization software*, ACM Trans. Math. Software, **7** (1981), No. 1, 17–41
[24] Pohlheim, H., *Examples of objective functions*, `http://www.geatbx.com` (2006)
[25] Scarf, H., *The approximation of fixed points of a continuous mapping*, SIAM Journal on Applied Mathematics, **15** (1967), No. 5, 1328–1343
[26] Scarf, H. and Hansen, T., *The computation of economic equilibria*, Yale University Press New Haven, Connecticut (1973)
[27] Thele, R. L., *Iterative techniques for approximation of fixed points of certain nonlinear mappings in Banach spaces*, Pacific J. Math., **53** (1974), 259–266
[28] Zhang, H. C. and Hager, W. W., A non monotone line search technique and its application to unconstrained optimization, SIAM J. Optim., **14** (2004), No. 4, 1043–1056

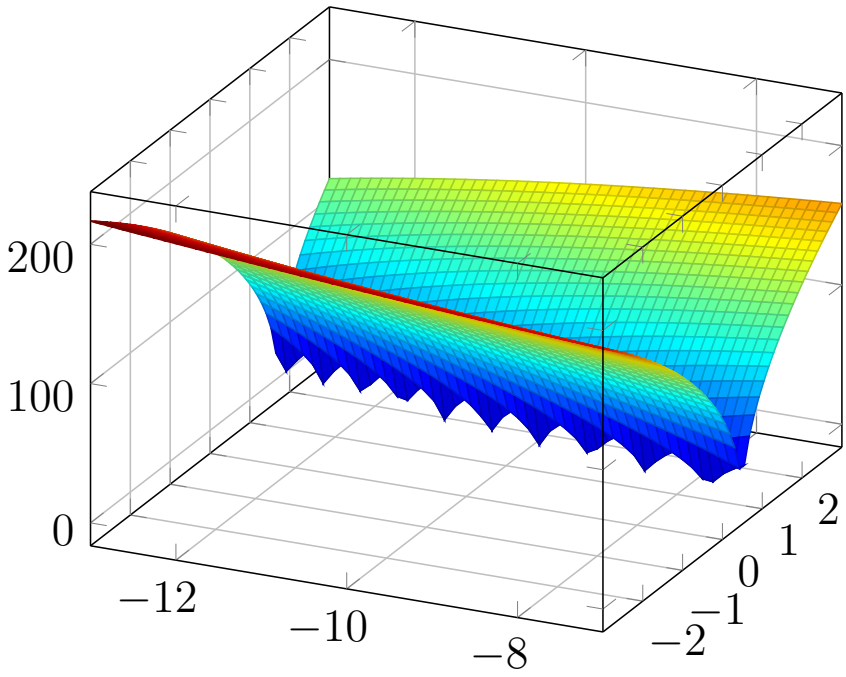

FIGURE 1. Bukin Function N. 6

x_0, y_0	-15,0	-10,0	-11,3	-13, -3
Newton iterations	∞	∞	∞	∞
BFGS iterations BFGS func evals BFGS gradient evals	∞	∞	∞	∞
CG iterations CG func evals CG gradient evals	∞	∞	∞	∞
NCG iterations NCG func evals NCG gradient evals NCG hessian evals	∞	∞	∞	∞
PLH iterations PLH label evals	199 308	35 64	92 146	104 158

Table 1. Comparing numerical results for Bukin Function N. 6

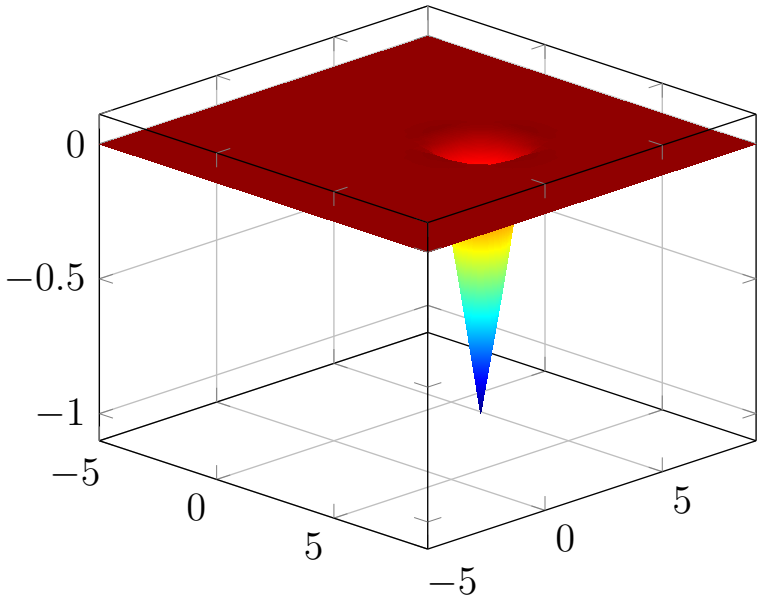

FIGURE 2. Easom Function

x_0, y_0	0,0	1,1	2,2	3,3
Newton iterations	$\neq$	$\neq$	$\neq$	4
BFGS iterations	∞	$\neq$	3	3
BFGS func evals			9	5
BFGS gradient evals			9	5
CG iterations	∞	$\neq$	1	2
CG func evals			18	5
CG gradient evals			6	5
NCG iterations	2	$\neq$	3	4
NCG func evals	14		9	5
NCG gradient evals	15		11	8
NCG hessian evals	2		3	4
PLH iterations	$\neq$	$\neq$	11	5
PLH label evals			17	11

Table 2. Comparing numerical results for Eason Function

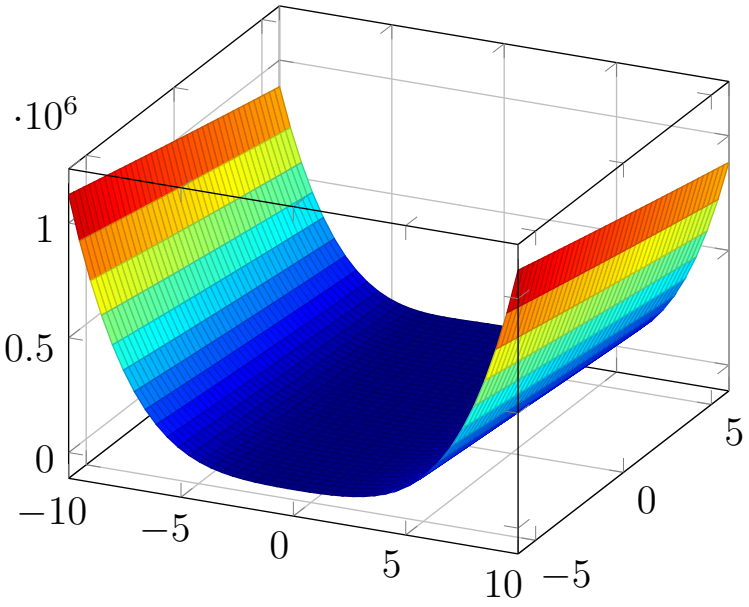

FIGURE 3. Rosenbrock Function

x_1, x_2	0,0	2,2	3,3	10,10	10000,10000
Newton iterations	4	7	8	9	∞
BFGS iterations	21	20	38	49	$\neq$
BFGS func evals	26	30	57	69	
BFGS gradient evals	26	30	57	69	
CG iterations	17	$\neq$	31	$\neq$	$\neq$
CG func evals	33		78		
CG gradient evals	33		78		
NCG iterations	33	24	31	51	$\neq$
NCG func evals	53	37	44	83	
NCG gradient evals	85	60	74	133	
NCG hessian evals	33	24	31	51	
PLH iterations	115	91	100	169	308
PLH label evals	174	136	145	214	380

Table 3. Comparing numerical results for Rosenbrock Function

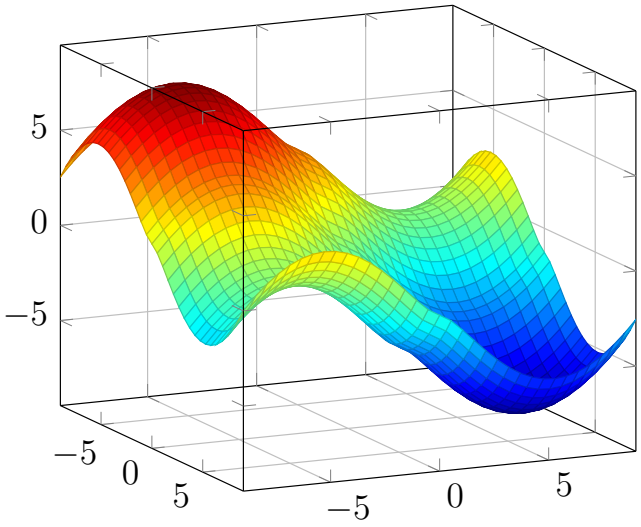

FIGURE 4. Schwefel Function for $n = 2$

x_0	c_1	c_2	c_3	c_4	c_5	c_6	c_7
Newton iterations	2	∞	∞	∞	4	∞	∞
BFGS iterations	3	$\neq$	5	$\neq$	4	$\neq$	∞
BFGS func evals	4		6		6		
BFGS gradient evals	4		6		6		
CG iterations	1	$\neq$	$\neq$	$\neq$	2	$\neq$	∞
CG func evals	15				6		
CG gradient evals	4				6		
NCG iterations	3	$\neq$	3	$\neq$	4	∞	∞
NCG func evals	4		5		5		
NCG gradient evals	6		7		8		
NCG hessian evals	3		3		4		
PLH iterations	14	20	43	44	45	50	45
PLH label evals	28	29	51	53	53	58	58

Table 4. Comparing numerical results for Schwefel Function ($n = 2$)

[1] DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
NORTH UNIVERSITY CENTER AT BAIA MARE, TECHNICAL UNIVERSITY OF CLUJ-NAPOCA
VICTORIEI 76, 430122 BAIA MARE, ROMANIA
E-mail address: andrei.bozantan@gmail.com, vberinde@ubm.ro, vasile_berinde@yahoo.com

[2] DEPARTMENT OF MATHEMATICS AND STATISTICS
KING FAHD UNIVERSITY OF PETROLEUM AND MINERALS
DHAHRAN, KINGDOM OF SAUDI ARABIA
E-mail address: vasile.berinde@gmail.com