Scholars Bulletin *(Engineering)*
An Official Publication of "Scholars Middle East Publishers"
Dubai, United Arab Emirates
Website: http://scholarsbulletin.com/
ISSN 2412-9771 (Print)
ISSN 2412-897X (Online)

The Compare of Two Different Non-Monotone Strategies for Solving the Derivative-Free Wedge Trust-Region Method

Fenghua Liu[1*], Qinghua Zhou[2]
[1]College of Computer Science and Engineering, Cangzhou Normal University, 060001, Hebei Province, China
[2]School of Applied Mathematics, Beijing Normal University, Zhuhai, China

***Corresponding author**
Fenghua Liu

Article History
Received: 02.12.2018
Accepted: 11.12.2018
Published: 30.12.2018

DOI:
10.21276/sb.2018.4.12.3

Abstract: In this paper, we compare the difference of non-monotone strategies to solve the wedge trust region method for derivative-free optimization. The wedge trust region method based on traditional trust region which is designed for derivative-free problems. Considering the effectiveness of the non-monotone strategy compared with monotone ones, we combined the non-monotone strategy into wedge trust region methods. The computational results show that the both two strategies have their respective advantages.
Keywords: non-monotone method, unconstrained optimization, derivative free optimization, wedge trust region.

INTRODUCTION

We consider the unconstrained optimization problem:

$$\min f(x),\ x \in R^n, \tag{1}$$

Where the objective function $f(x)$ is a smooth function from R^n to R, the $\nabla f(x)$ and $\nabla^2 f(x)$ are not available for any x.

In 1994, Powell [1] first proposed a derivative-free optimization method that used interpolation to approximate the objective and the constraints. The trust region interpolation models have the following form:

$$m_k(x_k+s) = f(x_k) + g_k^T s + \frac{1}{2} s^T G_k s, \tag{2}$$

Where $g_k \in R^n$ is a vector of R^n, G_k is a square symmetric matrix of dimension n. Since the gradient and Hessian matrix of the objective can not be calculated, we depart from many trust-region algorithms in that g_k and G_k will not be determined by the first and second derivative of $f(x)$, but rather by the function value at past points, that is we demand

$$m(y_k) = f(y_k) \tag{3}$$

for each vector y_k in a set $I = \{y^0, y^1, \ldots, y^{p-1}\}$. The cardinality of I must be equal to

$$p = \frac{1}{2}(n+1)(n+2). \tag{4}$$

The parameter p and the interpolation points set I must be poised with the purpose of ensuring the uniqueness and existence of the quadratic model [2-5]. When the model m_k is determined by the above conditions, the interpolation set is nonsingular.

The wedge trust region method is firstly proposed by Marazzi in his dissertation [6]. Firstly, the variable $y^{l_{out}}$ as the point to be replaced at the k-th iteration, i.e., $y^{l_{out}} = \arg\max_{y \in I} \|y - x_k\|$, the farthest one from the current iteration center x_k. Then, the "taboo region" T_k is defined in R^n (please refer to [7, 8], which contains all the points $x_k + s$

would result in a non-poised interpolation set if they are included in the interpolation set in place of $y^{l_{out}}$. A so-called "wedge" is defined that is a set W_k which contains T_k and that is designed to avoid points that are too near T_k. The wedge is added to the trust region sub-problem, which can be rewritten as follows:

$$\max_{0\le j\le m(k)} \{f_{k-j}\}, \qquad (5)$$

$$s.t.\|s\| \le \Delta_k \qquad (6)$$

$$s \notin W_k . \qquad (7)$$

In 1986, Grippo *et al.*, [9] proposed a non-monotone strategy, and the general non-monotone form is:

$$f_{l(k)} = f(x_{l(k)}) = \max_{0\le j\le m(k)} \{f_{k-j}\}, \quad k = 0,1,2,\ldots, \text{where } m_0 = 0$$

$0 \le m_k \le \min\{m_{k-1}+1, M\} (k \ge 1)$, and $M \ge 0$ is an integer. The sequence $\{f(x_k)\}$ is non-increasing. Since then, the non-monotone technique has been exploited by many researchers [10, 11] and a lot of numerical tests have showed that the non-monotone technique proposed by Grippo *et al.*, [9] is efficient.

In 2008, Gu and Mo [12] introduced another non-monotone strategy. They replaced $f_{l(k)}$ with

$$D_k = \begin{cases} f(x_k) & k = 1 \\ \eta_k D_{k-1} + (1-\eta_k) f(x_k) & k \ge 2 \end{cases}, \qquad (8)$$

This non-monotone technique is robust which is showed by numerical experiments in [12]

and the parameter $\eta_k = \dfrac{f(x_k)}{f(x_k) - D_{k-1}}$.

It is obvious that the non-monotone techniques may improve both the possibility of finding the global optimum and the rate of convergence. However, although the non-monotone technique has many advantages, it contains some drawbacks [13]. In order to overcome those disadvantages, Ahookhosh *et al.*, in [14, 15] proposed a new non-monotone technique. They define

$$R_k = \eta_k f_{l(k)} + (1-\eta_k) f_k , \qquad (9)$$

Where, $\eta_{\min} \in [0,1)$, $\eta_{\max} \in [\eta_{\min}, 1]$ and $\eta_k \in [\eta_{\min}, \eta_{\max}]$. This non-monotone technique is efficient and robust which is showed by numerical experiments in [15].

The rest of this paper is organized as follow. In section 2, the non-monotone strategies for wedge trust region will be established, and the algorithm analysis is interpreted. Numerical results are proved in section 3 which is indicated that the new methods have a lot of differences for unconstrained optimization problems. Finally, some conclusions are given in section 4.

The non-monotone wedge trust region algorithm

Step 0 Set the trial parameters, an initial trust region radius $\Delta_k > 0$, and an initial guess x_0. The interpolation set $Y_k = x_k \cup I$, $I = \{y^1, y^2, \cdots, y^m\}$, and it such that $f(x_k) \le f(y) \forall y \in I$.

Step 1 According to the current iteration point x_k, compute

$$y^{l_{out}} = \arg\max_{y\in I} \|y - x_k\|.$$

Step 2 Construction quadratic model m_k and define the wedge constraint W_k.

Step 3 Solve the sub-problem (2) and compute the trial step s_k, and calculate

$$r_k = \frac{Ared(d_k)}{\Pr ed(d_k)} = \frac{f(x_k) - f(x_k + s_k)}{m(0) - m_k(s_k)}, \ \overline{r_k} = \frac{R_k - f(x_k + s_k)}{m(0) - m_k(s_k)} \text{ or } \overline{r_k} = \frac{D_k - f(x_k + s_k)}{m(0) - m_k(s_k)}.$$

Step 4 Update the trust region radius Δ_k with the following:

$$\Delta_{k+1} = \begin{cases} \beta_4 \|s_k\|,_k r_1 < \alpha_1 \text{且} \overline{r}_k > \alpha_1; \\ \beta_1 \|s_k\|, r_1 < \alpha_1 \text{且} \overline{r}_k \le \alpha_1; \\ \Delta_k, \alpha_1 \le r_k < \alpha_2; \\ \beta_2 \Delta_k, \alpha_2 \le r_k \le \alpha_3; \\ \beta_3 \Delta_k, r_k > \alpha_3. \end{cases}$$

Step 5 Update the interpolation set and the iteration point, if it is a successful iteration, that is $\overline{r_k} > \alpha_1$, then $x_{k+1} = x_k + s$, $Y = \{x_k\} \cup I / \{y^{l_{out}}\}$.

Else it is a unsuccessful iteration, that is to say $\overline{r_k} < \alpha_1$, then $x_{k+1} = x_k$,

$$Y = \begin{cases} \{x_k + s\} \cup I / \{y^{l_{out}}\}, if \left\|y^{l_{out}} - x_k\right\| \ge \|(x_k + s) - x_k\| \\ Y, \text{otherwise} \end{cases}.$$

Step 6 $k = k + 1$, go to step 1.

NUMERICAL RESULTS

In this section, we compare the quadratic version of the $a\lg$ with the $c\lg$. The difference of $a\lg$ and $c\lg$ is the non-monotone strategy, the non-monotone strategy in $a\lg$ is R_k and the method in $c\lg$ is D_k. We set the rotating is $\pi / 600$ used in quadratic model. The initial value $\gamma = 0.4$ is allowed to change over the iteration. The MATLAB source code for wedge trust region algorithm is in [16]. Specific ally, we select 39 trial problems, which come from the CUTE. In the Table 1, the name of 39 test questions and results are given. We define n is the dimension of the objective function, and nf is the calculative times of an experimental function value. f is the optimal point and the $wed\ act$ represents the number of wedge constraints play a role. The final value of parameter γ which is a parameter used to control the space of “taboo region” is given in the last column when the algorithms stop. The parameters in our algorithms are taken as follow, $\alpha_1 = 0.01$, $\alpha_2 = 0.95$, $\alpha_3 = 1.05$, $\beta_1 = 0.5$, $\beta_2 = 2$, $\beta_3 = 1.01$, $\beta_4 = 0.8$.

According to the Table-1, we can obtain the following results. Firstly, comparing nf between the $a\lg$ and $c\lg$, we can see that the $c\lg$ is better than the $a\lg$. The numbers of wins of two algorithms are 2 and 1, respectively. However, the γ between the $a\lg$ and the $c\lg$ is different from the result of the nf, the wedge constraint of $a\lg$ is very more active than the wedge constraint of $c\lg$ although some numbers of function evaluations are the same. In addition, the result show that the $a\lg$ is better than the $c\lg$ when we must iterate many times for finding the optimal point. Secondly, we will give some examples to carefully describe the difference between $a\lg$ and $c\lg$. For example, to the problem “ALLINITU”, the numbers of nf wins of $c\lg$ and $a\lg$ are 109 and 167, but the γ of $a\lg$ is more active than $c\lg$. For example, to the test problem ”ARWHEAD”, the γ is the same as the problem “ALLINITU”, but the nf $c\lg$ of is more than $a\lg$. In spite of the f is very small in the $c\lg$, it would waste so much time when the initial value is very large.

Table-1: The computational results

n	p	nf		f		wed act		final γ	
		clg	alg	clg	alg	clg	alg	clg	alg
2	BROWNBS	45	40	9.89E+11	9.90E+11	5	5	1.40E-05	1.41E-05
2	HIMMELBB	154	159	1.57E-09	4.82E-05	19	9	3.75E-05	2.30E-03
2	HIMMELBH	65	154	-1.0000	-0.9967	25	6	1.85E-14	8.57E-04
2	BEALE	153	158	2.33E-09	1.96E-04	9	4	8.88E-06	4.37E-04
3	BOX3	34	33	3.03E-33	3.03E-33	15	13	1.58E-15	2.13E-16
2	BRKMCC	48	26	0.1690	0.1690	15	16	1.38E-15	3.87E-15
4	BROWNDEN	174	201	8.58E+04	5.64E+05	26	5	3.45E-17	3.30E-03
3	CUBE	233	148	2.38E-08	5.10E+00	10	1	1.20E-06	2.02E-02
2	EXPFIT	83	154	2.41E-01	5.07E-01	29	5	7.59E-16	2.15E-02
3	GULF	100	237	1.39E-02	5.91E+00	7	8	3.05E-04	2.04E-05
4	ALLINITU	109	167	5.74E+00	5.82E+00	38	2	6.85E-16	1.39E-02
10	CRAGGLVY	757	230	1.89E+00	1.90E+03	34	7	7.59E-16	7.50E-04
6	BIGGS6	543	157	1.05E-07	5.95E-01	20	3	1.68E-07	1.01E-02
10	BROWNAL	224	224	2.44E-01	2.41E-01	3	5	1.10E-03	1.48E-04
2	ENGVAL1	160	204	3.23E-04	4.80E-02	7	7	1.10E-03	2.60E-03
2	HAIRY	91	129	2.00E+01	4.65E+02	14	1	4.42E-11	6.80E-03
2	HIMMELBG	91	154	0.00E+00	5.40E-05	65	6	1.6E-154	8.71E-04
2	FREUROTH	87	174	4.90E+01	4.90E+01	26	7	1.36E-15	5.29E-05
5	GENHUMPS	220	226	1.05E-01	6.15E+04	9	8	9.38E-04	8.28E-05
3	HATFLDD	112	164	2.34E-06	3.69E-01	8	2	4.81E-06	5.40E-03
10	BRYBND	217	217	1.28E+00	2.42E+00	7	4	7.12E-04	1.90E-03
3	PFIT1LS	301	180	2.00E+02	2.36E+02	12	6	1.36E-05	1.70E-03
4	WOODS	409	175	1.23E-30	3.68E+02	23	2	1.11E-06	2.49E-02
5	OBSORNEA	206	176	8.11E-05	1.42E-01	13	6	1.43E-07	2.71E-05
6	EDENSCH	199	187	1.03E+02	1.05E+02	32	2	4.93E-06	7.10E-03
6	HEART6LS	378	182	4.22E-01	1.80E+02	9	3	1.69E-05	8.30E-03
2	CLIFF	38	162	2.00E-01	2.06E-01	10	3	9.60E-07	4.20E-03
2	DENSCHNA	163	167	2.30E-03	1.27E-08	4	6	1.20E-03	1.00E-05
2	BROWNBS1	45	56	9.89E+11	9.83E+11	5	5	1.40E-05	4.41E-06
8	ARGLINC	46	49	2.32E+01	2.32E+01	1	1	8.10E-02	1.40E-03
10	ARGLINB	118	118	1.00E+00	4.63E+01	12	10	1.79E-15	1.85E-13
2	BOX2	20	21	3.03E-33	3.03E-33	11	10	1.27E-13	1.68E-12
15	DIXMAANI	548	292	1.00E+00	3.09E+01	29	5	6.23E-05	4.60E-03
2	SINEVAL	245	258	1.53E-41	1.15E+00	18	7	2.80E-18	1.75E-04
15	ARWHEAD	501	315	2.19E-13	8.18E+00	19	4	3.01E-09	2.70E-03
2	SISSER	134	186	5.21E-08	3.08E-06	9	9	7.27E-04	2.20E-04
10	VARDIM	215	213	4.83E+03	2.40E+03	7	6	1.49E-04	3.93E-04
3	BARD	134	180	8.20E-03	4.72E-02	23	6	7.76E-17	5.40E-03
2	JENSMP	90	170	1.24E+02	1.30E+02	33	9	1.00E-15	1.37E-04

CONCLUSIONS

In this paper, we compared the difference of two non-monotone strategies to solve the wedge trust region method for derivative-free optimization. The results of numerical texts show that the two methods have their relative merits. In general, we can make a decision from the two strategies when the dimension of the initial point is different. In the near future we will learn and seek more new efficient non-monotone strategies for the wedge trust region methods to solve the derivative-free unconstrained optimization problem.

ACKNOWLEDGMENTS

This work is supported by the Teacher Research Capacity Promotion Program of Beijing Normal University, Zhuhai.

REFERENCES

1. Powell, M. J. (1994). A direct search optimization method that models the objective and constraint functions by linear interpolation. In *Advances in optimization and numerical analysis* (pp. 51-67). Springer, Dordrecht.
2. Conn, A. R., Scheinberg, K., & Vicente, L. N. (2008). Geometry of interpolation sets in derivative free optimization. *Mathematical programming*, *111*(1-2), 141-172.

3. Fasano, G., Morales, J. L., & Nocedal, J. (2009). On the geometry phase in model-based algorithms for derivative-free optimization. *Optimization Methods & Software*, *24*(1), 145-154.
4. Conn, A. R., Scheinberg, K., & Vicente, L. N. (2009). Global convergence of general derivative-free trust-region algorithms to first-and second-order critical points. *SIAM Journal on Optimization*, *20*(1), 387-415.
5. Moré, J. J., & Sorensen, D. C. (1983). Computing a trust region step. *SIAM Journal on Scientific and Statistical Computing*, *4*(3), 553-572.
6. Marazzi, M., & Nocedal, J. (2002). Wedge trust region methods for derivative free optimization. *Mathematical programming*, *91*(2), 289-305.
7. Powell, M. J. D. (2007). New developments of NEWUOA for minimization without derivatives. *Tech. Rep*.
8. Morales, J. L. (2007). *A trust region based algorithm for unconstrained derivative-free optimization*. tech. rep., Departamento de Matemáticas, ITAM.
9. Chamberlain, R. M., Powell, M. J. D., Lemarechal, C., & Pedersen, H. C. (1982). The watchdog technique for forcing convergence in algorithms for constrained optimization. In *Algorithms for Constrained Minimization of Smooth Nonlinear Functions* (pp. 1-17). Springer, Berlin, Heidelberg.
10. Grippo, L., Lampariello, F., & Lucidi, S. (1986). A nonmonotone line search technique for Newton's method. *SIAM Journal on Numerical Analysis*, *23*(4), 707-716.
11. Sun, W. (2004). Nonmonotone trust region method for solving optimization problems. *Applied Mathematics and Computation*, *156*(1), 159-174.
12. Gu, N. Z., & Mo, J. T. (2008). Incorporating nonmonotone strategies into the trust region method for unconstrained optimization. *Computers & Mathematics with Applications*, *55*(9), 2158-2172.
13. Zhang, H., & Hager, W. W. (2004). A nonmonotone line search technique and its application to unconstrained optimization. *SIAM journal on Optimization*, *14*(4), 1043-1056.
14. Ahookhosh, M., & Amini, K. (2012). An efficient nonmonotone trust-region method for unconstrained optimization. *Numerical Algorithms*, *59*(4), 523-540.
15. Ahookhosh, M., Amini, K., & Peyghami, M. R. (2012). A nonmonotone trust-region line search method for large-scale unconstrained optimization. *Applied Mathematical Modelling*, *36*(1), 478-487.
16. Bongartz, I., Conn, A. R., Gould, N., & Toint, P. L. (1995). CUTE: Constrained and unconstrained testing environment. *ACM Transactions on Mathematical Software (TOMS)*, *21*(1), 123-160.