Comparisons of two Non-Monotone Strategies for Solving Derivative-free Wedge Trust-Region Problems

Yan Li[1], Qinghua Zhou[1], Yan Zhan[2]
1School of Applied Mathematics, Beijing Normal University, Zhuhai, 519087, China
2Key Lab. of Machine Learning and Computational Intelligence, College of Mathematics and Information Science, Hebei University, Baoding, 071002, China

Abstract: In this paper, we compare the differences of non-monotone strategies to solve the wedge trust region method for derivative-free optimization. The non-monotone method is effective to resolve the trust region algorithm, and the wedge trust region method is projected for derivative-free problems. We combined the non-monotone strategy into wedge trust region methods, and the computational results showed that the two strategies have their respective advantages and disadvantages.

Keywords: Non-monotone method; Derivative free optimization, Wedge trust region

1. Introduction

In this paper, we consider the unconstrained optimization problem $\min f(x), \quad x \in R^n$ where the objective function $f(x)$ is a smooth function from R^n to R, $\nabla f(x)$ 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:

$$m_k(x_k+s)=f(x_k)+g_k^T s+(1/2)s^T G_k s,$$

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't be calculated, we demand $m(y_k)=f(y_k)$ for each vector y in a set $I=\{y^0, y^1, \mathbf{K}, y^{p-1}\}$. The cardinality of I must be equal to $p=(1/2)(n+1)(n+2)$ and 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, 3]. 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 [4]. Firstly, we define $y^{l_{out}}$ as the point to be replaced at the k-th iteration which is the farthest one from the current iteration center x_k. Then, we define the "taboo region" T_k [5, 6] in R^n. The wedge contain is added to the trust region sub-problem:

$$\min_s m_k(x_k+s)$$
$$s.t. \|s\| \le \Delta_k$$
$$s \notin W_k.$$

In 1986, Grippo et al. [7] proposed a non-monotone strategy, and the general non-monotone form is as $f_{l(k)}=f(x_{l(k)})=\max_{0\le j\le m(k)}\{f_{k-j}\}, \quad k=0,1,2,...,$ where $m_0=0$, $M \ge 0$ and $0 \le m_k \le \min\{m_{k-1}+1, M\}(k\ge 1)$ is an integer. The sequence $\{f(x_k)\}$ is non-increasing. Since then, the non-monotone technique has been exploited by many researchers [8, 9].

In 2008, Gu and Mo [10] introduced another non-monotone strategy. They replaced $f_{l(k)}$ with

$$h_k=\frac{f(x_k)}{f(x_k)-D_{k-1}}.$$

This non-monotone technique is robust which is showed by numerical experiments in [10, 11]. Ahookhosh et al. in [12] proposed a new non-monotone technique. They define $R_k=h_k f_{l(k)}+(1-h_k)f_k$, $h_{\min}\in[0,1), h_{\max}\in[h_{\min},1]$, This non-monotone technique is efficient and robust which is showed by numerical experiments in [12].

These two kinds of nonmonotone methods have their adv antes and disadvantages respectively in dealing with opti mization problems. In the present paper, we compar-e the wedge trust region method and two non-monotone techniques respectively. Our numerical results show that one can find a better method to solve the wedge trust region problems for using non-monotone strategy.

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. Some conclusions are given in section 4.

2. The Non-monotone Wedge Trust Region Algorithm

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

Step 2. According to the current iteration point x_k, compute $y^{l_{out}} = \arg\max_{y \in I} \|y - x_k\|$.

Step 3. Construction quadratic model m_k and define the wedge constraint W_k.

Step 4. Solve the sub-problem and compute the trial step s_k, and calculate

$$r_k = \frac{Ared(d_k)}{\Pr ed(d_k)} = \frac{f(x_k) - f(x_k + \mathrm{s}_k)}{m(0) - \mathrm{m}_k(\mathrm{s}_k)},$$

$$\overline{r_k} = \frac{R_k - f(x_k + \mathrm{s}_k)}{m(0) - \mathrm{m}_k(\mathrm{s}_k)}, \quad \overline{r_k} = \frac{D_k - f(x_k + \mathrm{s}_k)}{m(0) - \mathrm{m}_k(\mathrm{s}_k)}.$$

Step 5. Update the trust region radius Δ_k with the following:

$$\Delta_{k+1} = \begin{cases} b_4 \|s_k\|,_k r_1 < a_1 且 \overline{\mathrm{r}_k} > a_1; \\ b_1 \|s_k\|, r_1 < a_1 且 \overline{\mathrm{r}_k} \le a_1; \\ \Delta_k, a_1 \le r_k < a_2; \\ b_2 \Delta_k, a_2 \le r_k \le a_3; \\ b_3 \Delta_k, \mathrm{r}_k > a_3. \end{cases}$$

Step 6. Update the interpolation set and the iteration point, if it is a successful iteration, that is $\overline{r_k} > a_1$, 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 $\overline{r_k} < a_1$, 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 7. $k = k + 1$, go to step 2.

3. Numerical Results

In this section, we compare the quadratic version of the $a\lg$ with the $c\lg$. The non-monotone strategy in $a\lg$ is R_k and the method in $c\lg$ is D_k. The source code for algorithm is in [13], and we select 23 trial problems. In the Table 1, we define n is the dimension of the objective function, and nf is the calculative times, f is the optimal point, and the $wed\ act$ represents the number of wedge constraints play a role. The final value of parameter $g = 0.4$ is a parameter used to control the space of "taboo region". The parameters in our algorithms are taken as follow, $a_1 = 0.01$, $a_2 = 0.95$, $a_3 = 1.05$, $b_1 = 0.5$, $b_2 = 2$, $b_3 = 1.01$, $b_4 = 0.8$.

Table 1. The Computational Result (about nf and f)

n	p	nf		f	
		$c\lg$	$a\lg$	$c\lg$	$a\lg$
2	BROWNBS	45	40	9.89E+11	9.90E+11
10	HIMMELBB	154	159	1.57E-09	4.82E-05
2	HIMMELBH	65	154	-1.0000	-0.9967
2	BEALE	153	158	2.33E-09	1.96E-04
4	ALLINTU	109	167	5.74E+00	5.82E+00
6	BIGGS6	543	157	1.05E-07	5.95E-01
2	ENGVAL1	160	204	3.32E-04	4.80E-02
2	HAIRY	91	129	2.00E+01	4.65E+02
2	HIMMELBG	91	154	0.00E+00	5.40E-05
2	FREUROTH	87	174	4.90E+01	4.90E+01
5	GENHUMPS	220	226	1.05E-01	6.15E+04
3	HATFLDD	112	164	2.34E-06	3.69E-01
3	PFIT1LS	301	180	2.00E+02	2.36E+02
4	WOODS	409	175	1.23E-30	3.68E+02
5	OBSORNEA	206	176	8.11E-05	1.42E-01
6	HEART6LS	378	182	4.22E-01	1,80E+02
2	CLIFF	38	162	2.00E-01	2.06E-01
2	DENSCHNA	163	167	2.30E-03	1.27E-08
2	SINEVAI	245	258	1.53E-41	1.15E+00
15	ARWHEAD	501	315	2.19E-13	8.18E+00
2	SISSER	134	186	5.21E-08	3.08E-06
3	BARD	134	180	8.20E-03	4.72E-02
2	JENSMP	90	170	1.24E+02	1.30E+02

Table 2. The Computational Result (about $wed\ act$ and $final\ g$)

n	p	$wed\ act$		$final\ g$	
		$c\lg$	$a\lg$	$c\lg$	$a\lg$
2	BROWNBS	5	5	1.40E-05	1.41E-05
10	HIMMELBB	19	9	3.75-05	2.30E-03
2	HIMMELBH	25	6	1.85E-14	8.57E-04
2	BEALE	9	4	8.88E-06	4.37E-04
4	ALLINTU	38	2	6.85E-16	1.39E-02

6	BIGGS6	20	3	1.68E-07	1.01E-02
2	ENGVAL1	7	7	1.10E-03	2.60E-03
2	HAIRY	14	1	4.42E-11	6.80E-03
2	HIMMELBG	65	6	1.6E-154	8.71EE-04
2	FREUROTH	26	7	1.36E-15	5.29E-05
5	GENHUMPS	9	8	9.38E-04	8.28E-05
3	HATFLDD	8	2	4.81E-06	5.40E-03
3	PFIT1LS	12	6	1.36E-05	1.70E-03
4	WOODS	23	2	1.11E-06	2.49E-02
5	OBSORNEA	13	6	1.43E-07	2.71E-05
6	HEART6LS	9	3	1.69E-05	8.30E-03
2	CLIFF	10	3	9.60E-07	4.20E-03
2	DENSCHNA	29	5	6.23E-05	4.60E-03
2	SINEVAI	18	7	2.80E-18	1.75E-04
15	ARWHEAD	19	4	3.01E-09	2.70E-03
2	SISSER	9	9	7.27E-04	2.20E-04
3	BARD	23	6	7.76E-17	5.40E-03
2	JENSMP	33	9	1.00E-15	1.37E-04

According to the Table1 and 2: Firstly, we can see that $c\lg$ is better than $a\lg$ for nf, and the numbers of wins are 2 and 1. Respectively, for the wedge constraint g, $a\lg$ is very more active than $c\lg$ although some numbers of function evaluations are the same. In addition, the result show that the $a\lg$ is better than $c\lg$ when we must iterate many times for finding the optimal point. Secondly, we give some examples to carefully describe the difference between them. The numbers of nf wins are 109 and 167 in the problem "ALLINITU", but g of $a\lg$ is more active than $c\lg$. To the problem "ARWHEAD", the g is the same as the problem "ALLINITU", but the nf of $c\lg$ is more than the other. 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.

4 . 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 merit and demerit. In general, we can make a decision from the two strategies when the dimension of the initial is different. We will learn and seek more new efficient methods to solve the derivative-free unconstrained optimization problem.

Acknowledgements

This work is supported by the National Natural Science Foundations of China (61473111) and the Natural Science Foundation of Hebei University (No. 799207217069).

References

[1] M. J. D. Powell, A direct search optimization method that models the objective and constraint functions by linear interpolation. Advances in Optimization and Numerical Analysis, Proceedings of the Sixth Workshop on Optimization and Numerical Analysis, Oaxaca, Mexico, 1994, 275: 51-67.

[2] A.R. Coon, K. Scheinberg, and L. Vicente, Geometry of interpolation sets in derivative free optimization, Mathematical Programming, 2007, Series A, 111: 141-172.

[3] G. Fasano and J.L. Morales, On the geometry phase in model-based algorithms for derivative-free optimization, Opitimization Methods & Software, 2009, 24(1):145-154.

[4] M. Marazzi and J. Nocedal, Wedge trust region methods for derivative free optimization, Math. Program, 2002, Series A, 91: 289-305.

[5] M.J.D. Powell, New developments of NEWUOA for minimization without derivatives, Tech, Rep. DAMPT 2007/NA05, Department of Applied Mathematics and Theoretical physics, University of Cambridge, 2007.

[6] J.L. Morales: A trust region based algorithm for unconstrained derivative-free optimization, Tech. Rep, Departamento de Matemáticas, ITAM, 2007.

[7] R.M. Chamberlain and M.J.D. Powell: The watchdog technique for forcing convergence in algorithm for constrained optimization, Mathematical Programming Study 16, 1982, 1-17.

[8] L. Grippo, F. Lampariello and S. Lucidi, A nonmonotone line search technique for Newton's method, Society for Industrial and Applied Mathematics 23, 1986: 707-716.

[9] W.Y. Sun, Nonmonotone trust region method for solving optimization problems, Applied Mathematics and Computation 156, 2004:159-174.

[10] N.Z. Gu and J.T. Mo, Incorporating nonmonotone strategies into the trust region for unconstrained optimization, Computers and Mathematics with Applications 55, 2008: 2158–2172.

[11] H.C. Zhang and W.W. Hager, A nonmonotone line search technique and its application to unconstrained optimization, SIAM J. Optim., 2004, 14 (4): 1043-1056.

[12] M. Ahookhosh, K. Amini and M. R. Peyghami, A nonmonotone trust-region line search method for large-scale unconstrained optimization, Applied Mathematical Modelling 36, 2012: 478-487.

[13] Bongartz, A.R. Conn, N.I.M. Gould, and Ph. L. Tiont, CUTE: Constrained and unconstrained testing environment. ACM Transactions on Mathematical Software, 1995, 21(1): 123-160.