A New Non-monotone Wedge Trust-region Method for Derivative-free Unconstrained Optimization

Yan Li[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: A new wedge trust region method based on traditional trust region is designed for derivative-free problems. By combining a non-monotone strategy into wedge trust region methods, the new method is more efficient than the traditional one, and the results support the effectiveness of the new composite strategy.

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

1. Introduction

In this paper, we think about the unconstrained optimization problem: $\min f(x),\ \ x \in R^n$, where the function $f(x)$ is continuous and its derivatives cannot be explicitly computed [1, 2].

Considering the class of derivative-free trust-region methods, many algorithms can be found in the literature. They have global convergence quality [3, 4] for solving the following model, $m_k(x_k+s)=f(x_k)+g_k^T s+\frac{1}{2}s^T G_k s$

where $g_k \in R^n$ and the $n\times n$ symmetric matrix G_k are determined by the model interpolates f at a set of points $m(x_k)=f(x_k), m_k(y^l)=f(y^l), l=1,2,\ldots,m$ and $Y_k=\{y^1,y^2,\ldots,y^m\}\cup\{x_k\}$ is the interpolation point set. $m=n+1$ can be chosen when using the linear model, and $m=(1/2)(n+1)(n+2)$ should be chosen when using the quadratic model. Actually, this paper combined the linear and quadratic model for solving the derivative-free optimization effectively. The following is the construction method of the $(1/2)(n+1)(n+2)-(n+1)$ interpolation points when the linear model is transformed into the quadratic model [5, 6].

Suppose the present point is x_k, and the interpolation point set of linear model is $\{x_k\}$, where $k=0,1,2,\ldots,n$, then the quadratic model can be constructed. Set $x_k=y_0$, and the other points are respectively y, $k=1,2,\ldots,n$, then:

$$y_{n+1}=\begin{cases} y_0-q(y_i-y_0), s_i=-1 \\ y_0+2q(y_i-y_0), s_i=1 \end{cases}.$$

If $f(x_i)<f(x_0)$, then $s_i=1$, and else $s_i=-1$. Set $i(p,q)=2n+1+p+(1/2)(q-1)(q-2)$, $(1\le p<q\le n)$, and

$$y_{n+i(p,q)}=y_0+q\left[s_{n+p}(y_{n+p}-y_0)+s_{n+q}(y_{n+q}-y_0)\right].$$

Actually, the wedge trust region method is to compute a trial step s_k by solving

$$\begin{aligned} &\min_s m_k(x_k+s) \\ &s.t. \|s\|\le \Delta_k \\ &s\notin W_k. \end{aligned}$$

where W_k is a set which contains the “taboo region” area [7], and its purpose is to avoid the new point falling into it. The trail step s_k is calculated by the method which is introduced in [8]. We should set the $y^{l_{out}}$ which is the farthest satellite from the current iterate x_k, and it can guarantee the virtue of the models.

In 1982, the first non-monotone technique was proposed by Chamberlain et al. [9] for constrained optimization to overcome the Maratos effect. As it developed [10-13], Zhang et al. [14] found that non-monotone technique still had some drawbacks and they put forward a new non-monotone model

$$r_k=\frac{C_k-f(x_k+s_k)}{m_k(x_k)-m_k(x_k+s_k)},$$

$$C_k=\begin{cases} f(x_k), k=0, \\ \frac{h_{k-1}Q_{k-1}C_{k-1}+f(x_k)}{Q_k}, k\le 1, \end{cases}\quad Q_k=\begin{cases} 1, k=0, \\ h_{k-1}Q_{k-1}+1, k\ge 1, \end{cases}$$

where

$h_{\min}\in[0,1)$, $h_{\max}\in[h_{\min},1]$, $h_{k-1}\in[h_{\min},h_{\max}]$ are two given constants.

The rest of this paper is organized as follow. In section 2, the new non-monotone wedge trust region algorithm will be established, and the algorithm analysis is interpreted. Numerical results are proved in section 3 which indicate that the new method is very efficient for unconstrained optimization problems. Some conclusions are given in section 4.

2. A new Non-monotone Wedge Trust Region Algorithm

Step 1. Set the current point x_k, an initial trust region radius $\Delta_k > 0$, and the parameter of wedge trust region $g = 0.4$. The interpolation set is $Y = \{y^1, y^2, \mathbf{L}, y^m\}$, which satisfy $f(x_k) \le f(y)$, and $y^{l_{out}} = \arg\max_{y \in Y} \|y - x_k\|$.

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

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

$$r_k = \frac{Ared(d_k)}{\Pr ed(d_k)} = \frac{f(x_k) - f(x_k + s_k)}{m(0) - m_k(s_k)}, \quad \overline{r_k} = \frac{C_k - f(x_k + s_k)}{m(0) - m_k(s_k)}$$

Step 4. If $10 * eps \le \Delta_k < 0.1$ or $\overline{r_k} < 0.5$, set model as 'quadratic', go to step 1; else set model as 'linear'.

Step 5. Update the trust region radius with the following Algorithm analysis: if the model is the linear model, then $\Delta 1_{k+1}$; if the model is the quadratic model, then $\Delta 2_{k+1}$.

$$\Delta 1_{k+1} = \begin{cases} b_1\|s_k\|, & \overline{r_k} < a_1 \\ \Delta_k, & a_1 \le \overline{r_k} < a_2 \\ b_2\Delta_k, & a_2 \le \overline{r_k} < a_3 \\ b_3\Delta_k, & \overline{r_k} > a_3. \end{cases} \quad \Delta 2_{k+1} = \begin{cases} b_1\|s_k\|, & \overline{r_k} < 0, iter = 0 \\ b_3\Delta_k, & \overline{r_k} < 0, iter = 1 \\ b_2\Delta_k, & \overline{r_k} > 0, \|s_k\| = \Delta_k \\ \Delta_k, & \text{otherwise}. \end{cases}$$

$$a_1 = 0.01, a_2 = 0.95, a_3 = 1.05, \ b_1 = 0.5, b_2 = 2, b_3 = 1.01.$$

Step 6. if $f(x_k + s_k) < f(x_k)$, then $x_+ = x_k + s_k$, $Y_+ = \{x_k\} \cup Y_k \setminus \{y^{lout}\}$; else $x_{k+1} = x_k$,

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

Step 7. $x_k \leftarrow x_+, Y_k \leftarrow Y_+, \Delta_k \leftarrow \Delta_+$, go to step 2.

3. Numerical Results

In this section, we select 15 problems, which are from the CUTE [15]. In this work, the traditional wedge trust method $\lg 1$ and the new method $\lg 2$ are compared according to the number of function evaluations. In the following table, the name of 45 test questions and results are given. We define n as the dimension of the objective function, nf 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 g which is a parameter used to control the space of "taboo region" is given in the last column when the algorithms stop.

Table 1. Comparison Non-monotone Wedge Trust Region Algorithm with Wedge Trust Region (about nf and f)

n	problem	nf		f	
		lg1	lg2	lg1	lg2
6	BIGGS6	355	211	2.44E-01	3.79E-09
10	BROWNAL	344	221	9.03E+11	2.10E-28
2	HAIRY	52	192	3.90E+02	2.00E+01
2	HIMMELBG	99	151	1.43E-07	0.00E+00
2	FREUROYH	80	189	4.93E+01	4.90E+01
5	GENHUMPS	209	204	7.57E+03	3.79E-09
3	HATFLDD	112	181	6.02E-012	6.62E-08
10	BRYBND	260	223	3.64E+11	2.40E-29
3	PFIT1LS	222	171	2,51E+02	2.02E+02
5	OBSORNEA	279	121	1.77E-01	7.44E-05
6	EDENSCH	116	183	1.07E+02	1.03E+02
6	HEART6LS	1200	185	1.29E+02	3.16E-01
2	SISSER	236	147	8.72E-08	1.52E-58
3	BARD	94	156	7.84E-02	8.20E-03
2	JENSMP	66	148	2.65E+02	1.24E+02

Table 2. Comparison Non-monotone Wedge Trust Region Algorithm with Wedge Trust Region (about $wed\ act$ and g)

n	problem	$wed\ act$		g	
		lg1	lg2	lg1	lg2
6	BIGGS6	17	5	4.01E-07	6.10E-03
10	BROWNAL	46	2	6.54E-15	1.07E-01
2	HAIRY	14	2	2.24E-12	1.36E-02
2	HIMMELBG	53	6	6.0E-157	3.82E-05
2	FREUROYH	21	4	3.73E-15	9.49E-05
5	GENHUMPS	23	2	2.54E-17	5.00E-03
3	HATFLDD	20	4	2.29E-15	3.40E-03
10	BRYBND	31	8	1.81E-14	3.10E-03
3	PFIT1LS	11	9	5.14E-06	2.50E-03
5	OBSORNEA	17	9	9.04E-10	1.84E-04
6	EDENSCH	28	4	1.07E-15	6.10E-03
6	HEART6LS	20	6	1.51E-06	6.14E-02
2	SISSER	36	3	1.61E-15	7.10E-03
3	BARD	26	106	9.12E-16	3.70E-03
2	JENSMP	32	3	5.24E-15	4.50E-03

In this experiment, there are 14 problems of the new algorithm performing better than the traditional wedge trust region method considering the number of function value's calculations, and the final *g* of the new algorithms is bigger than the traditional methods for every problem. At the same time, the new method used less time getting the optimal solution. According to the experimental results, we can see that the new method is very effective and easy to be implemented.

4 . Conclusions

In this paper, we pretend a new non-monotone wedge trust region method for derivative free unconstrained optimization. The mixed non-monotone algorithm of linear and quadratic model is very efficient for obtaining the optimal function.

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] A.R. Coon, K. Scheinberg and Ph.L. Toint, "Recent progress in unconstrained nonlinear optimization without derivatives", Mathematical Programming,1997, 79:397-414.

[2] M.J.D. Powell, "A new algorithm for unconstrained optimization" [J], Nonlinear programming, 1970: 31-65.

[3] Y. X. Yuan and L.F. Niu, "A new trust-region algorithm for nonlinear constrained optimization". J. Comp. Math., 2010, 28, 72-86.

[4] M.J.D. Powell, "UOBYQA: Unconstrained optimization by quadratic approximation" [J], Mathematical Programming, 2002, 92: 555-582.

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

[6] M.J.D. Powell, "Direct search algorithms for constrained optimization calculations" [J], Mathematical Programming, 1998, 7: 187-336.

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

[8] Moré, J. J., and D.C.B. Soresen, "Computing a trust region step", SIAM Journal on Scientific and Statistical Computing, 1983, 4(3): 553-572.

[9] 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.

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

[11] N.Y. Deng, Y. Xiao and F.J. Zhou, "Nonmonotone trust region algorithm", Journal of Optimization Theory and Application, 1993, 76: 259-285.

[12] Ph. L. Toint, "Non-monotone trust-region algorithm for nonlinear optimization subject to convex constraints", Mathmatical Programming, 1997, 77: 69-94.

[13] Y. Yang and W.Y. Sun, "A new nonmonotone self-adaptive trust region algorithm with line search", Chinese Journal of Engineering Mathematics, 2007, 24(5): 788-794.

[14] H.C. Zhang and W.W. Hager, "A nonmonotone line search technique and its application to unconstrained optimization", SIAM J. Optim, 2014, 14 (4): 1043-1056.

[15] [15] M. Marazzi, "Software for derivative-free unconstrained nonlinear optimization", Information on http://www.eecs.northwestern.edu/~nocedal/wedge.html(2011).