南京航空航天大学

论文集

（二〇〇三年）第16册

理学院

南京航空航天大学科技部编

二〇〇四年三月

理 学 院

081 系

082 系

序号	姓名	职称	单位	论文题目	刊物、会议名称	年、卷、期
01	倪勤	正高	081	A globally convergent method of moving asymptotes with trust region technique	Optimigation Method and Software	2003 年 18 卷 3 期
02	倪勤	正高	081	浅谈德国高校数学教育	中国教育理论杂志	2003 年 16 期
03	戴彧虹 倪勤	副高 正高	外 081	Testing Different Conjugate Gradient Methods for Large-scale unconstrained optimization	Journal of Computational Mathematics	2003 年 21 卷 3 期
04	戴华	正高	081	对称正交对称矩阵反问题的最小二乘解	计算数学	2003 年 25 卷 1 期
05	戴华 姚承勇	正高 本科	081 081	Jacobi矩阵逆特征问题解存在的条件	高等学校计算数学学报	2003 年 25 卷 1 期
06	戴华	正高	081	一类对称正交反对称矩阵反问题	工程数学学报	2003 年 20 卷 3 期
07	解惠青 戴华	博士 正高	081 081	On the sensitivity analysis of multiple eigenvalues of nonsymmetric matrix pencil	Linear Algebra and Its Applications	2003 年 374 卷
08	古志鸣	正高	081	数学——不仅仅需要逻辑	大学数学	2003 年 19 卷 5 期
09	张鲁明	正高	081	A Conservative Numerical Scheme for a Class of Nonlinear *Schrodinger* Equation with Wave Operator	Applied Mathematics and Computation	2003 年 145 卷
10	刘心声 王金德	副高 正高	081 外	Testing for increasing convex order in several populations	The Annals of the Institute of Statistical Mathematics	2003 年 5 卷 1 期
11	盛松柏	正高	081	Optimality Conditions of a Class of Special Nonsmooth Programming	Journal of Computational Math.	2003 年 21 卷 6 期
12	李鹏同	正高	081	Ring isomorphisms and pentagon subspace lattices	Linear Algebra and its Applications	2003 年 367 卷
13	李鹏同	正高	081	Elementary operators on J-subspace lattice algebras	Bull.Austral.Math.Soc.	2003 年 68 卷
14	李鹏同	正高	081	Invertibility preserving linear maps on J-subspace lattice algebras	Linear Algebra and its Applications	2003 年 372 卷
15	安玉坤	正高	081	Periodic solutions of a Nonlinear Suspension Bridge Equation with Damping and Nonconstant Load	Journal of Mathematical Analysis and Application	2003 年 279 卷 2 期
16	安玉坤	正高	081	On the Coupled System of second and fourth order elliptic	Applied Mathematics and Computation	2003 年 140 卷 2-3 期

				equations		
17	安玉坤	正高	081	吊桥耦合系统静态解的存在唯一性	西北师范大学学报（自然科学版）	2003 年 39 卷 1 期
18	岳勤	副高	081	域上矩阵的相似标准型	徐州师范大学学报（自然科学版）	2003 年 21 卷 4 期
19	岳勤	副高	081	On Lifting Projective Modules Modulo an Essential Socle	Communications in Algebra	2003 年 31 卷 9 期
20	耿显明	副高	081	宏观经济均衡发展的数学理论（III）	应用数学与计算数学学报	2003 年 17 卷 1 期
21	顾玉娣 戴振东	副高 正高	081 051	具有线性回归特征的多组实验结果对比分析方法的研究	南京航空航天大学学报	2003 年 35 卷 5 期
22	顾玉娣	副高	081	大学生数学建模竞赛活动的认识与实践	南京航空航天大学学报（社会科学版）	2003 年 5 卷
23	王正盛	中级	081	阻尼振动系统中的逆二次特征值问题	2003 年第七届全国计算数学年会	2003 年
24	钟宝江	中级	081	On the restrarted Arnoldi method for large nonsymmetric matrix problems	2003 年第七届全国计算数学年会	2003 年
25	钟宝江	中级	081	GMRES 方法的收敛率	高等学校计算数学学报	2003 年 25 卷 3 期
26	钟宝江	中级	081	如何促进学生有效理解和接受数学的抽象问题	南京航空航天大学学报（社会科学版）	2003 年 5 卷
27	李艳 叶尔骅	初 级 正高	081 081	有优先权开关寿命连续型两部件冷贮备可修系统的可靠性分析	兰州大学学报	2003 年 39 卷 2 期
28	李秀娟	中级	081	求解多目标优化问题的随机梯度遗传算法	南京航空航天大学学报	2003 年 35 卷 4 期
29	杨秀绘	中级	081	一个退化抛物方程解的极大值点的注记	南京大学学报数学半年刊	2003 年 20 卷 1 期
30	丁洁玉 叶尔骅 杨纪龙	硕士 正高 副高	081 081 081	对数正态与正态分布定时截尾寿命试验参数的近似置信域	南京航空航天大学学报	2003 年 35 卷 6 期
31	许克祥	初级	081	Coloring of some integer distance graphs	Journal of Southeast University (English Editon)	2003 年 19 卷 4 期
32	施大宁	正高	082	A modified mean field theory for spin system with orbital degeneracy	Chinese Physics letters	2003 年 20 卷 5 期
33	赵志敏 陈玉明 俞晓磊	正高 硕士 本科	082 082 082	血液的荧光光谱特性分析及应用研究	光谱学与光谱分析	2003 年 25 卷 5 期
34	俞晓磊 赵志敏	本科 正高	082 082	应用光激励 SMA 技术实现结构损伤控制	量子电子学	2003 年 20 卷 2 期
35	廖立	副高	082	开设文科大学物理 强化学生	南京航空航天大学学	2003 年 5 卷

				素质教育	报（社会科学版）	23 期
36	王开圣	中级	082	轧辊堆焊用全纤维电热处理炉	机械科学研究院哈尔滨焊接研究所、中国机械工程学会焊接学会	2003 年 10 期
37	王开圣 刘小廷	中级 副高	082 082	十一线电位差计定标的可操作性	大学物理实验	2003 年 16 卷 2 期
38	张永梅	中级	082	Evolution of Spin and Charge in a System with Interacting Impurity and Conductance Electrons	Chinese Physics Letters	2003 年 20 卷 11 期
39	罗建	硕士	082	Slave-boson approach to electronic Raman spectra in high-Tc cuprates	Physica C	2003 年 399 卷

Optimization Methods and Software, Vol. 18, No. 3, June 2003, pp. 283–297

A GLOBALLY CONVERGENT METHOD OF MOVING ASYMPTOTES WITH TRUST REGION TECHNIQUE

QIN NI*

Nanjing University of Aeronautics and Astronautics, 210016 Nanjing, P.R. China

(Received 6 September 2001; Revised 29 January 2003; In final form 29 January 2003)

The method of moving asymptotes is known to work well for certain problems arising in structural optimization. A globally convergent method of moving asymptotes with trust region technique is proposed in this paper. A convex separable subproblem is solved in each iteration. The choice of asymptotes is controlled by the trust region radius such that global convergence of the algorithm is obtained. In addition, preliminary numerical tests are given.

Keywords: Method of moving asymptotes; Trust region method; Structural optimization

1 INTRODUCTION

We consider algorithms which combine moving asymptotes and trust region methods and which are designed for solving the following bound constrained optimization problem

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & x^l \le x \le x^u, \end{array} \tag{1.1}$$

where $x \in R^n$, the function f is assumed to be continuous in Ω and twice continuously differentiable in the interior of Ω, $\Omega = \{x \in R^n\colon x^l \le x \le x^u\}$, x^l and x^u are finite fixed vectors. It is assumed that

$$b_m = \min_{1 \le i \le n} (x_i^u - x_i^l) \ge 0. \tag{1.2}$$

The method of moving asymptotes (MMA) was firstly presented in [7], and frequently used for structural optimization. Afterwards this method was further studied and developed. Zillober combined the method with a line search, and obtained a kind of efficient global convergent sequential convex programming method [11,12]. For other extensions and developments see Refs. [2] and [8]. However, the properties of the moving asymptotes approximation is so far not thoroughly studied, the choice of the asymptotes is not very reasonable, and a global technique is complicated [9]. These may be a critical point of MMA. It is hoped that the famous trust region technique can improve MMA.

* Corresponding author. E-mail: niqfs@mail.nuaa.edu.cn

ISSN 1055-6788 print; ISSN 1029-4937 online
DOI: 10.1080/1055678031000118491

The aim of this paper is to propose a globally convergent algorithm which combines the moving asymptotes approximation with the trust region technique. This algorithm is used to solve the bound constrained optimization problem. Through the convergence analysis of the algorithm the properties of moving asymptotes approximation and choice of asymptotes are analyzed and studied.

At first, we give some definitions and assumptions. The projected gradient $\nabla_\Omega f$ of f is defined by

$$[\nabla_\Omega f(x)]_i = \begin{cases} \dfrac{\partial f}{\partial x_i}, & x_i^l < x_i < x_i^u \\ \min\left(0, \dfrac{\partial f}{\partial x_i}\right), & x_i = x_i^l \\ \max\left(0, \dfrac{\partial f}{\partial x_i}\right), & x_i = x_i^u \end{cases} \tag{1.3}$$

A point $x \in \Omega$ is called to be a stationary point of problem (1.1) if $\nabla_\Omega f(x) = 0$.

The projection operator $\mathcal{P}$ on Ω is $\mathcal{P}(x) = (P_1[x_1], \ldots, P_n[x_n])^T$, where

$$P_i[x_i] = \begin{cases} x_i^l, & \text{if } x_i \le x_i^l \\ x_i^u, & \text{if } x_i \ge x_i^u \\ x_i, & \text{otherwise} \end{cases} \tag{1.4}$$

We denote the set of all active bound constraints by $I(x)$,

$$I(x) = \{i\colon x_i = x_i^u \text{ or } x_i = x_i^l,\ x \in \Omega\}.$$

This paper is organized as follows. We discuss the trust region subproblem in Section 2 and propose new algorithm in the subsequent section. The convergence of the algorithm is presented in Section 4 and numerical tests are given in Section 5.

2 TRUST REGION SUBPROBLEM

The algorithm for solving (1.1) is of trust region type. At each iterate, a subproblem is defined by

$$\begin{aligned} \min_{d\in R^n} \quad & m(x,d) = f(x) + \sum_{i=1}^{n} \phi_i(d_i) \\ \text{s.t.} \quad & d_i^l \le d_i \le d_i^u, \quad i = 1,\ldots,n \end{aligned} \tag{2.1}$$

where

$$d_i^u = \min(\Delta_i, x_i^u - x_i), \quad d_i^l = \max(-\Delta_i, x_i^l - x_i), \tag{2.2}$$

Δ_i is trust region radius,

$$\phi_i(d_i) = \begin{cases} \dfrac{g_i d_i (u_i - x_i)}{u_i - x_i - d_i} + \dfrac{\varepsilon_i d_i^2}{(u_i - x_i - d_i)(x_i - l_i + d_i)}, & i \in I_+ \\ \dfrac{g_i d_i (x_i - l_i)}{x_i - l_i + d_i} + \dfrac{\varepsilon_i d_i^2}{(u_i - x_i - d_i)(x_i - l_i + d_i)}, & i \in I_- \end{cases}, \tag{2.3}$$

$I_+ = \{i: g_i \geq 0\}$, $I_- = \{1, \ldots, n\}/I_+$, $(g_1, \ldots, g_n)^T = \nabla f(x)$. ε_i, l_i and u_i are parameters, which satisfy

$$l_i < x_i < u_i, \quad \varepsilon_i > 0, \quad u_i - x_i - d_i > 0, \quad x_i - l_i + d_i > 0, \tag{2.4}$$

for $d_i \in [d_i^l, d_i^u]$, $i = 1, \ldots, n$. Their detailed choice is discussed in Lemma 2.1.

For convenience, some notations are introduced:

$$a_i = u_i - x_i, \quad b_i = x_i - l_i, \quad v_i = u_i - l_i, \quad i = 1, \ldots, n. \tag{2.5}$$

It follows from (2.4) that

$$a_i > 0, \quad b_i > 0, \quad v_i > 0, \quad v_i = a_i + b_i. \tag{2.6}$$

It is easy to see that the subproblem (2.1) is separable and convex. It is equivalent to n independent one-dimensional bound constrained subproblems

$$\begin{aligned} \min_{d_i} \quad & \phi_i(d_i) \\ \text{s.t.} \quad & d_i^l \leq d_i \leq d_i^u, \end{aligned} \tag{2.7}$$

$i = 1, \ldots, n$. Because $\phi_i(d_i)$ is a strictly convex function, there exists a unique optimal solution in (2.7). With some calculation, we have

$$\frac{\partial m(x,d)}{\partial d_i} = \phi_i'(d_i) = \begin{cases} \dfrac{g_i a_i^2}{(a_i - d_i)^2} + \dfrac{\varepsilon_i}{v_i}\left(\dfrac{a_i^2}{(a_i - d_i)^2} - \dfrac{b_i^2}{(b_i + d_i)^2}\right), & i \in I_+ \\ \dfrac{g_i b_i^2}{(b_i + d_i)^2} + \dfrac{\varepsilon_i}{v_i}\left(\dfrac{a_i^2}{(a_i - d_i)^2} - \dfrac{b_i^2}{(b_i + d_i)^2}\right), & i \in I_- \end{cases}, \tag{2.8}$$

where we observe that

$$\frac{\varepsilon_i d_i^2}{(u_i - x_i - d_i)(x_i - l_i + d_i)} = \frac{\varepsilon_i d_i^2}{(a_i - d_i)(b_i + d_i)} = \frac{\varepsilon_i}{v_i}\left(\frac{a_i^2}{a_i - d_i} + \frac{b_i^2}{b_i + d_i}\right) - \varepsilon_i.$$

The function $m(x, d)$ is a first-order approximation of f, i.e.

$$m(x, 0) = f(x), \quad \nabla_d m(x, 0) = (\phi_1'(0), \ldots, \phi_n'(0))^T = (g_1, \ldots, g_n)^T = \nabla f(x).$$

It is noted that an additional term in $\phi_i(d_i)$ is similar to those in [8] and [11]. However, it is somewhat different. The minimal solution point of $\phi_i(d_i)$ is bounded when ε_i approaches to 0_+, which is discussed later. Other MMA approximations with usual additional terms do not possess this property.

Let $\phi_i'(d_i) = 0$, i.e.

$$0 = \frac{g_i a_i^2}{(a_i - d_i)^2} + \frac{\varepsilon_i}{v_i}\left(\frac{a_i^2}{(a_i - d_i)^2} - \frac{b_i^2}{(b_i + d_i)^2}\right) \quad \text{for } i \in I_+$$

which means

$$\left(\frac{a_i - d_i}{b_i + d_i}\right)^2 = \frac{a_i^2}{b_i^2}\left(1 + \frac{v_i g_i}{\varepsilon_i}\right).$$

By (2.4) and (2.6), we obtain

$$\frac{a_i - d_i}{b_i + d_i} = \frac{a_i \lambda_i}{b_i}$$

where

$$\lambda_i = \left(1 + \frac{v_i g_i}{\varepsilon_i}\right)^{1/2}$$

and

$$d_i = \frac{a_i b_i (1 - \lambda_i)}{b_i + a_i \lambda_i} = \frac{a_i b_i (1 - \lambda_i^2)}{(1 + \lambda_i)(b_i + a_i \lambda_i)} = -\frac{a_i b_i v_i}{\varepsilon_i (1 + \lambda_i)(b_i + a_i \lambda_i)} g_i \quad \text{for } i \in I_+.$$

With the same deduction, we obtain the solution d_i' of $\phi_i'(d_i) = 0$ as

$$d_i' = -\alpha_i' g_i \tag{2.9}$$

where

$$\alpha_i' = \begin{cases} \dfrac{a_i b_i v_i}{\varepsilon_i(\lambda_i + 1)(b_i + \lambda_i a_i)}, & i \in I_+ \\ \dfrac{a_i b_i v_i}{\varepsilon_i(\lambda_i + 1)(a_i + \lambda_i b_i)}, & i \in I_- \end{cases} \qquad \lambda_i = \begin{cases} \left(1 + \dfrac{v_i g_i}{\varepsilon_i}\right)^{1/2}, & i \in I_+ \\ \left(1 - \dfrac{v_i g_i}{\varepsilon_i}\right)^{1/2}, & i \in I_- \end{cases}. \tag{2.10}$$

Hence the solution of (2.7) is described by

$$d_i = -\alpha_i g_i \tag{2.11}$$

where

$$\alpha_i = \begin{cases} \alpha_i', & \text{if } d_i' \in [d_i^l, d_i^u] \\ \dfrac{d_i^l}{-g_i}, & \text{if } d_i' < d_i^l \\ \dfrac{d_i^u}{-g_i}, & \text{if } d_i' > d_i^u \end{cases}. \tag{2.12}$$

It follows from (2.9) and (2.10) that if $g_i = 0$, then $d_i' = 0$; if $g_i \neq 0$, then

$$\lim_{\varepsilon_i \to 0_+} \varepsilon_i(\lambda_i + 1) = 0, \qquad \lim_{\varepsilon_i \to 0_+} \varepsilon_i(\lambda_i + 1)\lambda_i = \begin{cases} v_i g_i, & i \in I_+ \\ -v_i g_i, & i \in I_- \end{cases}$$

and

$$\lim_{\varepsilon_i \to 0_+} d_i' = \begin{cases} -b_i, & i \in I_+ \\ a_i, & i \in I_- \end{cases}$$

which means the boundedness of the minimal solution of $\phi_i(d_i)$.

We define three sets by

$$J_1 = \{i : d_i' \in [d_i^l, d_i^u]\}, \quad J_2 = \{i : d_i' < d_i^l\}, \quad J_3 = \{i : d_i' > d_i^u\}. \tag{2.13}$$

In the following lemma, some conditions for parameters l_i, u_i and ε_i are formulated, which guarantee the boundedness of α_i'.

LEMMA 2.1 *Assume that the parameters l_i, u_i and ε_i satisfy the following conditions:*

$$\begin{cases} \max\{c_0(x_i^u - x_i^l) + \Delta_i,\ c_1(x_i^u - x_i^l)\} \le x_i - l_i \le c_2(x_i^u - x_i^l) \\ \max\{c_0(x_i^u - x_i^l) + \Delta_i,\ c_1(x_i^u - x_i^l)\} \le u_i - x_i \le c_2(x_i^u - x_i^l) \end{cases} \tag{2.14}$$

$$\max\left(\varepsilon_l, \frac{|g_i|}{b_\varepsilon}\right) \le \varepsilon_i \le \varepsilon_u \tag{2.15}$$

$i = 1, \dots, n$, *where* $0 < c_0 < c_1 < 1 < c_2$, $0 < \varepsilon_l < \varepsilon_u$, $b_\varepsilon > 0$.

Then, α_i' is bounded, i.e.

$$\alpha_l \le \alpha_i' \le \alpha_u, \quad i = 1, \dots, n \tag{2.16}$$

where

$$\alpha_l = \frac{c_1^2 b_m^2}{2\varepsilon_u(1 + 2c_2 b_\varepsilon b_M)} > 0, \quad \alpha_u = \frac{c_2^2 b_M^2}{2\varepsilon_l}, \quad b_M = \max_{1 \le i \le n}(x_i^u - x_i^l).$$

Proof From (2.10), (2.14) and (2.15), we obtain

$$1 \le \lambda_i^2 \le 1 + 2c_2 b_\varepsilon b_M$$
$$c_1^2 b_m^2 v_i \le a_i b_i v_i \le c_2^2 b_M^2 v_i$$

Hence, from $\lambda_i \ge 1$ and (2.6) it follows that

$$\varepsilon_i(\lambda_i + 1)(b_i + \lambda_i a_i) \ge \varepsilon_i(\lambda_i + 1)v_i \ge 2\varepsilon_l v_i$$

and

$$\varepsilon_i(\lambda_i + 1)(b_i + \lambda_i a_i) \le 2\varepsilon_i\lambda_i(b_i + \lambda_i a_i) \le 2\varepsilon_i\lambda_i^2 v_i \le 2\varepsilon_u(1 + 2c_2 b_\varepsilon b_M)v_i.$$

With the same deduction, we have

$$2\varepsilon_i v_i \le \varepsilon_i(\lambda_i + 1)(a_i + \lambda_i b_i) \le 2\varepsilon_u(1 + 2c_2 b_\varepsilon b_M)v_i.$$

These imply that

$$\alpha_l \le \alpha_i' \le \alpha_u, \quad i = 1, \dots, n.$$

$\alpha_l > 0$ because of (1.2). ■

It is noted that the condition (2.14) is similar to that in [9], however the left side inequalities are different due to the trust region constraint. This means that for given Δ_i, the parameters l_i and u_i have to satisfy

$$l_i < x_i - \Delta_i, \quad u_i > x_i + \Delta_i \tag{2.17}$$

which may be a reasonable control of l_i and u_i. In a practical implementation, Δ_i in (2.14) can be replaced by $\tilde{c}\Delta_i$, ($\tilde{c} \in (0, 1]$) when Δ_i is relatively large. For the convenience in the proof, we only consider the case $\tilde{c} = 1$.

In order to obtain a main results about the subproblem, we estimate the difference between $\phi_i(0)$ and $\phi_i(d_i)$ in the following lemma.

LEMMA 2.2 *Assume that d_i is the solution of (2.7) given in (2.11), then we have*

$$\phi_i(0)-\phi_i(d_i)=\begin{cases}\alpha_i'\beta_i g_i^2, & i\in J_1\\ -d_i^l\beta_i g_i, & i\in J_2\\ -d_i^u\beta_i g_i, & i\in J_3\end{cases} \tag{2.18}$$

where

$$\beta_i=\begin{cases}\dfrac{a_i\lambda_i}{(a_i+\alpha_i'g_i)(\lambda_i+1)}, & i\in I_+\cap J_1\\ \dfrac{b_i\lambda_i}{(b_i-\alpha_i'g_i)(\lambda_i+1)}, & i\in I_-\cap J_1\\ \dfrac{a_ib_i}{(a_i-d_i^l)(b_i+d_i^l)(\lambda_i+1)}, & i\in J_2\\ \dfrac{a_ib_i}{(a_i-d_i^u)(b_i+d_i^u)(\lambda_i+1)}, & i\in J_3\end{cases} \tag{2.19}$$

Proof If $i\in J_1$, then it follows from (2.3) and (2.10) that

$$\begin{aligned}\phi_i(0)-\phi_i(d_i)&=\phi_i(0)-\phi_i(-\alpha_i'g_i)\\ &=\alpha_i'g_i^2\beta_i.\end{aligned}$$

If $i\in J_2$, then $g_i>0,\ i\in I_+$ and

$$\begin{aligned}\phi_i(0)-\phi_i(d_i)&=\phi_i(0)-\phi_i(d_i^l)\\ &=\frac{-d_i^l}{a_i-d_i^l}\left(g_ia_i+\frac{\varepsilon_id_i^l}{b_i+d_i^l}\right)\\ &=\frac{-d_i^l}{(a_i-d_i^l)(b_i+d_i^l)}(g_ia_ib_i+d_i^l(g_ia_i+\varepsilon_i))\end{aligned}$$

The definition of J_2 means that $d_i^l>-\alpha_i'g_i$. Taking into account the condition (2.6), we obtain

$$\begin{aligned}\phi_i(0)-\phi_i(d_i)&\geq\frac{-d_i^l}{(a_i-d_i^l)(b_i+d_i^l)}(g_ia_ib_i-\alpha_i'g_i(g_ia_i+\varepsilon_i))\\ &=\frac{-d_i^lg_ia_ib_i}{(a_i-d_i^l)(b_i+d_i^l)}\left(1-\frac{v_i(a_ig_i/\varepsilon_i+1)}{(\lambda_i+1)(b_i+a_i\lambda_i)}\right)\quad\text{cf. (2.10)}\\ &=\frac{-d_i^lg_ia_ib_i}{(a_i-d_i^l)(b_i+d_i^l)}\frac{(\lambda_i+1)(b_i+a_i\lambda_i)-(v_i+a_i(\lambda_i^2-1))}{(\lambda_i+1)(b_i+a_i\lambda_i)}\quad\text{cf. (2.10)}\\ &=\frac{-d_i^lg_ia_ib_i}{(a_i-d_i^l)(b_i+d_i^l)(\lambda_i+1)}\frac{\lambda_iv_i}{b_i+a_i\lambda_i}\quad\text{cf. (2.6)}\\ &\geq\frac{-d_i^lg_ia_ib_i}{(a_i-d_i^l)(b_i+d_i^l)(\lambda_i+1)}\\ &=\beta_i(-d_i^l)g_i.\end{aligned} \tag{2.20}$$

If $i \in J_3$, then $-d_i^u > \alpha_i' g_i$, $g_i < 0$, $i \in I_-$ and

$$\begin{aligned}\phi_i(0) - \phi_i(d_i) &= \phi_i(0) - \phi_i(d_i^u) \\ &= \frac{d_i^u}{(a_i - d_i^u)(b_i + d_i^u)}((-g_i)a_i b_i - d_i^u((-g_i)b_i + \varepsilon_i))\end{aligned}$$

From the definition of J_3 and condition (2.6), we obtain

$$\begin{aligned}\phi_i(0) - \phi_i(d_i) &\geq \frac{d_i^u(-g_i)}{(a_i - d_i^u)(b_i + d_i^u)}(a_i b_i - \alpha_i'((-g_i)b_i + \varepsilon_i)) \\ &= \frac{d_i^u(-g_i)a_i b_i}{(a_i - d_i^u)(b_i + d_i^u)} \frac{\lambda_i v_i}{(\lambda_i + 1)(a_i + \lambda_i b_i)} \\ &\geq \frac{d_i^u(-g_i)a_i b_i}{(a_i - d_i^u)(b_i + d_i^u)(\lambda_i + 1)} \\ &= \beta_i d_i^u(-g_i). \end{aligned} \tag{2.21}$$

Thus, the lemma is proved. ■

Now we define

$$h(x) = P[x - D\nabla f(x)] - x \tag{2.22}$$

where $D = \mathrm{diag}(\alpha_1', \ldots, \alpha_n')$, $\alpha_1', \ldots, \alpha_n'$ are the same as in (2.10), then a main result is described in the following theorem.

THEOREM 2.3 *Assume that the conditions* (2.14) *and* (2.15) *are satisfied, then we have*

$$f(x) - m(x,d) \geq \beta \min\left(1, \frac{\Delta m}{b_M}\right) \|h(x)\|^2, \tag{2.23}$$

where $\beta = \min_{1 \leq i \leq n}(\beta_i/\alpha_i') > 0$, $\Delta_m = \min_{1 \leq i \leq n} \Delta_i$.

Proof If $i \in J_1$, then

$$h_i(x) = P_i[x_i - \alpha_i' g_i] - x_i = -\alpha_i' g_i.$$

From (2.18), we obtain

$$\phi_i(0) - \phi_i(d_i) = \frac{\beta_i}{\alpha_i'} h_i^2(x). \tag{2.24}$$

If $i \in J_2$, then we have from (2.13) that $d_i^l > -\alpha_i' g_i$, $g_i > 0$ and $x_i - \alpha_i' g_i < x_i^u$ because of $\alpha' > 0$ and $d_i^l \leq 0$. Consider three cases in J_2.

(1) If $x_i = x_i^l$, then

$$h_i(x) = P_i[x_i - \alpha_i' g_i] - x_i = 0.$$

From (2.3) and $d_i^l = \max\{-\Delta_i, x_i^l - x_i\} = 0$, it follows that

$$\phi_i(0) - \phi_i(d_i) = \phi_i(0) - \phi_i(d_i^l) = 0. \tag{2.25}$$

(2) If $x_i > x_i^l$, $x_i - x_i^l \leq \Delta_i$, then $d_i^l = x_i^l - x_i$, $x_i - \alpha_i' g_i < x_i + d_i^l = x_i^l$, and

$$h_i(x) = P_i[x_i - \alpha_i' g_i] - x_i = x_i^l - x_i = d_i^l.$$

From (2.20), we have that

$$\phi_i(0) - \phi_i(d_i) = \frac{\beta_i}{\alpha_i'} |d_i^l| |\alpha_i' g_i| \geq \frac{\beta_i}{\alpha_i'} h_i^2(x). \tag{2.26}$$

(3) If $x_i > x_i^l$, $x_i - x_i^l > \Delta_i$, then $d_i^l = -\Delta_i$. From (2.22), we obtain

$$|h_i(x)| = |P_i[x_i - \alpha_i' g_i] - x_i| \leq \min(x_i - x_i^l, |\alpha_i' g_i|).$$

From (2.20), we obtain

$$\phi_i(0) - \phi_i(d_i) \geq -d_i^l g_i \beta_i \geq \frac{\Delta_i \beta_i}{\alpha_i'} (\alpha_i' g_i) \frac{x_i - x_i^l}{x_i^u - x_i^l} \geq \frac{\beta_i}{\alpha_i} \frac{\Delta_i}{x_i^u - x_i^l} h_i^2(x). \tag{2.27}$$

If $i \in J_3$, then we have that $g_i < 0$ and $x_i - \alpha_i' g_i > x_i^l$ because of $\alpha' > 0$ and $d_i^u \geq 0$. With a similar deduction, we obtain

$$\begin{aligned} \phi_i(0) - \phi_i(d_i) &= \frac{\beta_i}{\alpha_i'} h_i^2(x), \quad \text{if } x_i = x_i^u, \\ \phi_i(0) - \phi_i(d_i) &\geq \frac{\beta_i}{\alpha_i'} h_i^2(x), \quad \text{if } x_i > x_i^u, x_i^u - x_i \leq \Delta_i. \end{aligned} \tag{2.28}$$

If $x_i < x_i^u$, $x_i^u - x_i > \Delta_i$, then $d_i^u = \Delta_i$, and

$$\phi_i(0) - \phi_i(d_i) \geq \frac{\beta_i}{\alpha_i'} \frac{\Delta_i}{x_i^u - x_i^l} h_i^2(x). \tag{2.29}$$

The inequalities (2.24)–(2.29) together imply that (2.23) is valid. In addition, Lemma 2.1 and (2.19) imply that $\beta > 0$. ■

In addition, we have the following statement.

LEMMA 2.4 *Let $x \in \Omega$, D is defined by* (2.22), l_i, u_i, ε_i *satisfy* (2.14) *and* (2.15). *Then x is a stationary point for* (1.1) *if and only if*

$$h(x) = P[x - tD\nabla f(x)] - x = 0$$

for all $t \geq 0$.

Proof From Lemma 2.1, we obtain that D is diagonal and positive definite. Hence, this lemma results immediately from Proposition 1.35 of [1]. ■

3 ALGORITHM

We are now able to outline the new algorithm, which combines the moving asymptotes approximation with trust region technique.

ALGORITHM 3.1

Step 0 *Choose* $x_0 \in \Omega$, $k = 0$. *Choose the constants* $c_0, c_1, c_2, \varepsilon_l, \varepsilon_u, b_\varepsilon, \eta_1, \eta_2, r_1, r_2$ *and* r_3 *such that* $0 < c_0 < c_1 < 1 < c_2, 0 < \varepsilon_l < \varepsilon_u, b_\varepsilon > 0, 0 \le \eta_1 < \eta_2 < 1, 0 < r_1 \le r_2 < 1 < r_3$, *and define the initial trust region radius* $\Delta_i^0 > 0, i = 1, \dots, n$.

Step 1 *Compute* $f(x^k)$, g_k, *and define a subproblem (2.1). Let* d^k *be its solution* (*see* (2.7)). *If* $d^k = 0$, *then stop.*

Step 2 *Compute*

$$\rho_k = \frac{f(x^k + d^k) - f(x^k)}{m(x^k, d^k) - f(x^k)} \tag{3.1}$$

If

$$\rho_k \ge \eta_1, \tag{3.2}$$

then $x^{k+l} = x^k + d^k$, *otherwise* $x^{k+1} = x^k$.

Step 3 *Update* Δ_k.

$$\Delta_i^{k+1} \in \begin{cases} [r_2\Delta_i^k, \Delta_i^k], & \text{if } \rho_k < \eta_2 \\ [\Delta_i^k, r_3\Delta_i^k], & \text{otherwise} \end{cases} \tag{3.3}$$

$i = 1, \dots, n$.

Step 4 *Update* l_i^{k+1}, u_i^{k+1} *and* ε_i^{k+1} *such that* (2.14) *and* (2.15) *are satisfied.*

Step 5 $k = k + 1$, *go to Step 1.*

Remarks

(1) An iteration in Algorithm 3.1 is called successful if the test (3.2) is satisfied, and unsuccessful otherwise.

(2) If the solution of (2.7), d (see (2.9) and (2.10)) is nonzero, then the test (3.2) is easily satisfied, if the trust region radius is small enough. If $d = 0$, then a stationary point is reached, which is described in the following lemma.

LEMMA 3.1 *Let* d *be a solution of* (2.7). *Then* $d = 0$ *if and only if* $\nabla_\Omega f(x) = 0$.

Proof From (2.9) and $\alpha_i' \ge \alpha_l > 0$ in Lemma 2.1, it follows that

$$d_i = 0 \iff g_i = 0, \quad \text{for all } i \in J_1.$$

If $i \in J_2$, then from (2.11) to (2.13) we have $d_i = d_i^l$, $g_i > 0$, $\min(0, g_i) = 0$, because $\alpha_i' > 0$ and $d_i^l \le 0$ (see (2.2)). Hence, we obtain

$$d_i = 0 \iff d_i^l = 0 \iff x_i^l = x_i \text{ (because } \Delta_i > 0),$$

i.e.

$$d_i = 0 \iff x_i^l = x_i, \text{ and } \min(0, g_i) = 0 \quad \text{for all } i \in J_2 \iff (\nabla_\Omega f(x))_i = 0.$$

With a similar deduction, we obtain

$$d_i = 0 \iff x_i^u = x_i, \text{ and } \max(0, g_i) = 0 \quad \text{for all } i \in J_3 \iff (\nabla_\Omega f(x))_i = 0. \quad \blacksquare$$

This lemma means that if x is not a stationary point, then d in (2.9) is nonzero.

4 GLOBAL CONVERGENCE PROPERTY OF THE ALGORITHM

In this section we discuss the global convergence of Algorithm 3.1. We first define

$$S = \{k\colon \text{the } k\text{th iteration is successful in Algorithm 3.1}\}$$

and the curvature of the differentiable function f along the step d ($d \neq 0$) and based at the point $x \in \Omega$ by

$$w(f, x, d) = \frac{1}{\|d\|^2}[f(x+d) - f(x) - \nabla f(x)^T d]. \tag{4.1}$$

In a similar way we define

$$\tilde{w}(m, x, d) = \frac{1}{\|d\|^2}[m(x, d) - m(x) - \nabla_d m(x, 0)^T d] \tag{4.2}$$

where $d \neq 0$. It is noted that the function f is assumed to be continuous in Ω and twice continuously differentiable in the interior of Ω in the following lemmas and theorems. Now we have the statement for $w(f, x, d)$ and $\tilde{w}(m, x^k, d^k)$.

LEMMA 4.1 *There exist constants $c_3 > 0$ and $c_4 > 0$ such that*

$$|w(f, x, d)| \leq c_3, \tag{4.3}$$

$$\|\nabla f(x)\|_\infty \leq c_4, \tag{4.4}$$

where $d \in R^n$, $d \neq 0$, x, $x + d \in \Omega$.

The proof of this lemma is easy and is omitted.

LEMMA 4.2 *There exists a constant $c_5 > 0$ such that*

$$0 < \tilde{w}(m, x^k, d^k) \leq c_5 \tag{4.5}$$

where $d^k \neq 0$ is the solution of (2.7) at x^k,

$$c_5 = \frac{c_4 c_0 b_m + \varepsilon_u}{c_0^2 b_m^2}. \tag{4.6}$$

Proof From (4.1) and (2.1) it follows that

$$\begin{aligned} \|d^k\|^2 \tilde{w}(m, x^k, d^k) &= (m(x^k, d^k) - m(x^k, 0) - \nabla_d m(x^k, 0)^T d^k) \\ &= \sum_{i=1}^n \phi(d_i^k) - \sum_{i=1}^n \phi'(0) d_i^k \\ &= \sum_{i=1}^n (\phi(d_i^k) - g_i d_i^k) \\ &= \frac{1}{2} \sum_{i=1}^n \alpha_{ki} (d_i^k)^2 \end{aligned} \tag{4.7}$$

where α_{ki} is obtained after simple calculation

$$\alpha_{ki} = \begin{cases} \dfrac{2g_i}{a_i^k - d_i^k} + \dfrac{2\varepsilon_i}{(a_i^k - d_i^k)(b_i^k + d_i^k)}, & i \in I_+ \\ \dfrac{-2g_i}{b_i^k + d_i^k} + \dfrac{2\varepsilon_i}{(a_i^k - d_i^k)(b_i^k + d_i^k)}, & i \in I_- \end{cases}.$$

From (2.14), (2.15) and (4.4), it follows that

$$a_i^k - d_i^k \geq c_0(x_i^u - x_i^l) \geq c_0 b_m, \quad b_i^k + d_i^k \geq c_0 b_m,$$

and

$$0 < \alpha_{ki} \leq 2c_5.$$

■

If we define

$$\Delta_{km} = \min_{1\leq i\leq n}(\Delta_i^k), \quad \Delta_{kM} = \max_{1\leq i\leq n}(\Delta_i^k), \quad c_6 = \left(\frac{\max_{1\leq i\leq n}(\Delta_i^0)}{\min_{1\leq i\leq n}(\Delta_i^0)}\right)^2$$

then it is easy to see from Algorithm 3.1 that

$$\Delta_{kM}^2 \leq c_6 \Delta_{km}^2. \tag{4.8}$$

The following statement is somewhat different from general trust region methods, but the proof is similar.

LEMMA 4.3 *Consider a sequence $\{x^k\}$ generated by Algorithm 3.1, and assume that there exists a constant $\varepsilon > 0$ such that*

$$\|h(x^k)\| \geq \varepsilon \tag{4.9}$$

for all k, where $h(x^k)$ is defined by (2.22). Then there is a constant $c_7 > 0$ such that

$$\Delta_{km} \geq c_7 \tag{4.10}$$

Proof From (4.8), lemmas 4.1 and 4.2, it follows that

$$\begin{aligned} |f(x^k + d^k) - m(x^k, d^k)| &\leq \frac{1}{2}\|d^k\|^2 |w(f, x^k, d^k) - \tilde{w}(m, x^k, d^k)| \\ &\leq \frac{1}{2}\|d^k\|^2 (|w(f, x^k, d^k)| + |\tilde{w}(m, x^k, d^k)|) \\ &\leq \frac{1}{2}\|d^k\|^2 (c_3 + c_5) \leq \frac{1}{2}(c_3 + c_5)\Delta_{kM}^2 \\ &\leq \frac{1}{2}(c_3 + c_5) c_6 \Delta_{km}^2 \end{aligned} \tag{4.11}$$

We assume, without loss of generality, that

$$\varepsilon < \min\left\{1, b_M \left(\frac{(c_3 + c_5)c_6}{2(1 - \eta_2)\beta}\right)^{1/2}\right\}, \tag{4.12}$$

where β refers to Theorem 2.3, η_2 refers to Step 3 in Algorithm 3.1. In order to derive a contradiction, we assume that there exists a k such that

$$\frac{1}{2}(c_3 + c_5)c_6 b_M \Delta_{km} \le r_1 \beta (1 - \eta_2)\varepsilon^2 \tag{4.13}$$

where r_1 refers to Step 0 in Algorithm 3.1 and r is defined as the first iteration number such that (4.13) holds.

With Step 3 in Algorithm 3.1 we obtain

$$\begin{aligned} c_5 c_6 b_M \Delta_{r-1,m} &\le (c_3 + c_5)c_6 b_M \Delta_{r-1,m} \\ &\le \frac{(c_3 + c_5)c_6 b_M \Delta_{rm}}{r_1} \\ &\le 2(1-\eta_2)\beta\varepsilon^2. \end{aligned} \tag{4.14}$$

From Theorem 2.3, it follows that

$$f(x_{r-1}) - m(x_{r-1} + d_{r-1}) \ge \beta \min\left(1, \frac{\Delta_{r-1,m}}{b_M}\right) \|h(x_{r-1})\|^2.$$

Without loss of generality, assume that

$$\frac{\Delta_{km}}{b_M} \le 1, \quad k = 1, 2, \ldots,$$

Then we have

$$f(x_{r-1}) - m(x_{r-1} + d_{r-1}) \ge \frac{\beta \Delta_{r-1,m}\varepsilon^2}{b_M}. \tag{4.15}$$

(3.1), (4.11), (4.14) and (4.15) together means that

$$\begin{aligned} |\rho_{r-1} - 1| &= \frac{|f(x_{r-1} + d_{r-1}) - m(x_{r-1} + d_{r-1})|}{|f(x_{r-1}) - m(x_{r-1} + d_{r-1})|} \\ &\le \frac{(c_3 + c_5)c_6 b_M \Delta_{r-1,m}}{2\beta\varepsilon^2} \\ &\le 1 - \eta_2. \end{aligned} \tag{4.16}$$

This implies that $\rho_{r-1} \ge \eta_2$, and therefore $\Delta_{rm} \ge \Delta_{r-1,m}$. Hence, we obtain that

$$\frac{1}{2}(c_3 + c_5)c_6 b_M \Delta_{r-1,m} \le \frac{1}{2}(c_3 + c_5)c_6 b_M \Delta_{rm} \le r_1 \beta (1 - \eta_2)\varepsilon^2 \tag{4.17}$$

which contradicts the assumption that r was the first index satisfying (4.13). Therefore, (4.13) never holds, and

$$\frac{1}{2}(c_3 + c_5)c_6 b_M \Delta_{km} > r_1 \beta (1 - \eta_2)\varepsilon^2$$

for all k. Hence we obtain

$$\Delta_{km} \ge c_7,$$

for all k, where

$$c_7 = \frac{2 r_1 \beta (1 - \eta_2)\varepsilon^2}{(c_3 + c_5)c_6 b_M}. \qquad \blacksquare$$