



求解一般界约束优化问题的积极集信赖域方法

王晓

中国科学院数学与系统科学研究院计算数学与科学工程计算研究所, 科学与工程计算国家重点实验室, 北京 100190
E-mail: wangxiao@lsec.cc.ac.cn

收稿日期: 2010-04-12; 接受日期: 2011-03-31

国家自然科学基金 (批准号: 10831006) 和中国科学院知识创新工程 (批准号: kjcx-yw-s7-03) 资助项目

摘要 本文提出了一种求解一般界约束优化问题的新方法. 每步迭代分为两个阶段. 在第一阶段, 从当前迭代点 x_k 出发, 沿着经过仿射变换后的梯度步, 得到试探点 x_k^1 , 记录下它的积极集. 这里用到的仿射变换矩阵不仅依赖于变量到边界的距离, 还依赖于当前迭代点的梯度以及该步迭代中的信赖域半径. 在第二阶段, 从 x_k^1 出发, 通过在积极约束的零空间里面求解一个信赖域子问题得到新的试探点. 然后判断是否接受这个试探点作为下一个迭代点. 文中证明了算法的全局收敛性, 并且迭代点列的每个聚点都是一阶稳定点. 文中还对国际著名的 CUTEr 算例库中所有的界约束优化问题进行了测试. 数值结果表明我们的方法是有效的, 并且可以与 L-BFGS-B 方法相媲美.

关键词 界约束优化 积极集 仿射变换 信赖域

MSC (2000) 主题分类 65K03

1 引言

本文研究的是具有如下形式的一般界约束优化问题:

$$\min f(x) \quad (1.1a)$$

$$\text{s.t. } l \leq x \leq u, \quad (1.1b)$$

这里 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 是二次连续可微函数,

$$l = (l^{(1)}, \dots, l^{(n)})^T, \quad u = (u^{(1)}, \dots, u^{(n)})^T.$$

同时可以允许对应某个变量的上下界是无穷, 也即对某个 i 可以允许 $l^{(i)} = -\infty$ 或者 $u^{(i)} = \infty$. 不失一般性, 我们总是假定对所有变量都有 $l^{(i)} < u^{(i)}$.

一般界约束优化问题在现实世界中广泛出现 (参见文献 [1]), 所以设计求解此类问题的有效算法显得尤为重要. 到目前为止, 求解 (1.1) 的方法主要分为两类, 积极集法和内点法.

积极集方法通常是先确定积极集, 然后为了提高收敛速度, 往往需要在积极集的零空间里面进行进一步搜索 (见文献 [2, 3]). 上述这些文章中, 作者都引入了信赖域技巧. 首先, 沿着投影梯度步通过在信赖域里极小化目标函数 f 的二次近似 ψ 得到一个试探步. 由此试探步确定出积极集后, 保证这些积极的变量不变, 通过改变那些非积极的变量来进一步减小 ψ . 文献 [1] 与上述两篇有所不同. 在第一阶段中, 他们执行的是非单调的投影梯度步, 然后在满足某些条件之下, 转入无约束优化阶段.

英文引用格式: Wang X. An active set trust region method for general bound constrained optimization (in Chinese). Sci Sin Math, 2011, 41(4): 377–391, doi: 10.1360/012010-266

仿射变换方法最早由 Dikin^[4] 提出, 用于求解线性规划问题. 之后这种方法又分别由 Barnes^[5] 和 Vanderber 等^[6] 独立提出. 在 1980 年, 用于求解凸二次规划问题的二阶仿射变换方法由 Dikin^[7] 引入. 在文献 [7] 中, 作者构造了一个以当前迭代点为中心的椭球, 它的轴长与能包含在可行域里的最大椭球的轴长成一个常数比例 $\beta \in (0, 1)$. 通过在该椭球上极小化目标函数来得到下一个迭代点. 对于应用仿射变换方法来求解二次规划问题, 不管是凸的还是非凸的, 已经有很多学者进行了研究, 例如文献 [8–11]. 之后, 用于求解带有线性约束的凸规划问题的仿射变换方法也被提出, 见文献 [12, 13]. 其中文献 [13] 针对一类特殊的凸函数提出了一种仿射变换方法, 这类函数的二阶 Hesse 阵满足一些不变性质. 文中在无任何非退化假设的情况下给出了全局收敛性的证明. 依照文献 [7] 中的方式定义了一个椭球, 然后在该椭球上极小化目标函数的二次近似, 得到试探步. 通过计算

$$x_{k+1} = x_k + s_k / \kappa$$

来产生下一个迭代点, 这里 $\kappa \geq 1$ 是依赖于目标函数的二次曲率的一个常数. 与文献 [7] 不同, 文献 [12] 针对带有线性约束的凸 (或者凹) 规划问题提出了一种仿射变换方法. 其中 β_k 不再是一个常数, 而是可以自适应变化的一个参数, 它的更新策略与经典的信赖域方法中信赖域半径的更新方式类似. 文献 [14] 基于一个等价的仿射变换下的 KKT 系统提出了一种新的仿射变换方法, 与前面的方法不同, 这种方法允许椭球信赖域与可行域相交.

信赖域方法是求解非线性规划问题的一类非常重要的方法 (见文献 [15–17]). 这类方法以当前点为中心定义了一个所谓的信赖域, 同时定义原问题的二次近似模型, 计算该模型在信赖域中 (我们称之为子问题) 的近似极小作为试探步. 在某些接受准则之下, 接受该试探步从而产生下一个迭代点. 在实际计算中, 信赖域的大小要根据算法在前面的表现进行适时地调整, 具体的调整策略可见文献 [18]. 通常情况下, 信赖域半径的调整依赖于基于目标函数的真实下降量与基于二次模型的预测下降量之间的比值 (称之为下降比). 信赖域技巧已经被广泛应用于求解各种优化问题. 另外, 信赖域的形状也不局限于标准的球形区域. Moré 在文献 [19] 总结了针对无约束优化问题的信赖域方法在算法与软件方面的发展, 并且集中阐述了不同形状的信赖域在对变量进行尺度变换和预优方面的重要性.

本文提出了一种求解 (1.1) 的新的积极集信赖域方法. 与传统的积极集方法有所不同, 我们在确定积极集阶段走的是一个经过仿射变换的梯度步, 而并非文献 [1–3] 中所采用的梯度步. 事实上, 我们之所以采用上述的策略, 是因为考虑到, 由于在界约束优化中, 迭代点总是限定在约束里面, 所以我们希望能够把边界的信息和梯度信息同时考虑, 从而确定出试探步. 这里定义的仿射变换矩阵不仅依赖于变量到边界的距离, 同时还依赖于当前迭代点的梯度以及信赖域半径. 而此处考虑信赖域半径, 是利用信赖域自适应变化的特点, 希望在算法收敛性和有效性方面有所帮助. 这种仿射变换矩阵的定义方式与文献 [14] 中的策略是不同的. 文献 [14] 中的定义并没有考虑到信赖域半径可能带来的影响. 执行仿射变换下的梯度步后, 记录下当前状态的积极集. 为了加快收敛速度, 考虑在积极集的零空间中进一步极小化二次模型. 在这个过程中, 为了保证可行, 我们把子问题定义在一个完全包含在可行域里的信赖域中, 这种策略在经典的仿射变换方法中常常被用到. 文中证明了这个新的算法是全局收敛的, 且每个聚点都是一阶稳定点. 文中还针对国际上广泛使用的算例库 CUTEr 中的所有界约束优化问题进行了测试, 实验结果表明我们的方法与著名的 L-BFGS-B 方法相比是非常有效的.

本文的架构如下. 第 2 节详细阐述了新的算法, 并在第 3 节中给出了算法的全局收敛性证明. 第 4 节对 CUTEr 算例库做了数值实验并对数值表现进行了讨论. 最后一节是文章的总结.

符号 在下面的内容中, 不加下标的范数符号 $\|\cdot\|$ 表示的是二范数. 定义 $g(x) = \nabla f(x)$. 上标 “ (i) ” 表示向量的分量. 下标 “ k ” 则表示迭代指标, 简写为 $f_k = f(x_k)$ 及 $g_k = g(x_k)$.

2 算法

本节详细阐述求解界约束优化问题 (1.1) 的一种新的积极集信赖域方法. 定义可行域

$$\mathcal{F} := \{x \in \mathbb{R}^n : l \leq x \leq u\}.$$

问题 (1.1) 的一阶最优性条件可以很容易地写出, 见文献 [20]. 假设 x_* 是一个一阶稳定点, 则如下的一阶必要性条件在 x_* 处成立:

$$\begin{cases} g^{(i)}(x_*) = 0, & \text{若 } l^{(i)} < x_*^{(i)} < u^{(i)}, \\ g^{(i)}(x_*) \geq 0, & \text{若 } x_*^{(i)} = l^{(i)}, \\ g^{(i)}(x_*) \leq 0, & \text{若 } x_*^{(i)} = u^{(i)}, \\ l \leq x_* \leq u. \end{cases} \quad (2.1)$$

这里我们引入一个新的向量 $v(x)$, 它的元素定义为

$$v^{(i)}(x) = \begin{cases} u^{(i)} - x^{(i)}, & \text{若 } g^{(i)}(x) \leq 0, \\ x^{(i)} - l^{(i)}, & \text{若 } g^{(i)}(x) > 0. \end{cases} \quad (2.2)$$

假设 $x_k \in \mathcal{F}$ 是算法产生的第 k 个迭代点. 记 s_k 是在第 k 步迭代中产生的试探步. 下面我们说明 s_k 及下一个迭代点 x_{k+1} 是如何产生的. 目标函数 $f(x)$ 在当前迭代点 x_k 处的二次近似为

$$\psi_k(x) \triangleq f_k + g_k^T(x - x_k) + \frac{1}{2}(x - x_k)^T B_k(x - x_k),$$

其中 $B_k = (\approx) \nabla^2 f(x_k)$. 类似于传统的积极集方法, 我们首先确定一个搜索方向, 然后沿着这个方向极小化 $\psi_k(x_k + d)$, 同时还要保证搜索方向的可行性. 这里我们选取的搜索方向是经过尺度变换的梯度步, 之所以采用这种方式, 是要把边界的信息也考虑进去. 按照如下方式定义仿射变换的对角矩阵 D_k :

$$(D_k)_{ii} = \min\{v_k^{(i)}, \Delta_k\}, \quad (2.3)$$

其中 Δ_k 表示当前迭代中信赖域半径. 搜索方向定义为

$$d_k(\tau) = -\tau \frac{D_k^2 g_k}{\|D_k g_k\|} \quad (\tau \geq 0). \quad (2.4)$$

下面我们解释定义 (2.3) 的优势. 一方面, 信赖域方法有很好的全局收敛性质, 之所以引入信赖域半径, 是希望保证我们的算法是全局收敛的, 这一点将在第 3 节中给予详细的说明. 另一方面, 当某个分量没有界约束要求或是允许它的界很宽松时, 相应地, $v_k^{(i)}$ 将会是无穷或是很大, 而引入 Δ_k 之后, 它的平衡作用将会使这种情况在一定程度上得到避免; 否则为了保证可行性, 一旦我们沿着 (2.4) 定义的方向进行搜索得到的步长将会很小, 这就会导致算法的低效.

沿着 (2.4) 定义的方向 $d_k(\tau)$ 搜索, 得到试探步 s_k^1 , 它满足

$$s_k^1 = \arg \min_d \{\psi_k(x_k + d) : d = d_k(\tau), \|d\| \leq \Delta_k, x_k + d \in \mathcal{F}\}. \quad (2.5)$$

事实上, (2.5) 是一个一维优化问题, 所以可以很容易地求解. 令

$$x_k^1 = x_k + s_k^1. \quad (2.6)$$

x_k^1 的某些变量可能位于边界上, 我们把这些积极的分量所对应的指标定义为积极集

$$I_k^1 = \{i : (x_k^1)^{(i)} = l^{(i)}\}, \quad I_k^2 = \{i : (x_k^1)^{(i)} = u^{(i)}\}, \quad (2.7)$$

并记

$$I = \{1, 2, \dots, n\}. \quad (2.8)$$

由于到目前为止, 我们只是在某一特定方向上极小化二次模型从而得到了 x_k^1 , 所以为了提高算法的收敛速度, 在得到 x_k^1 后, 我们并不直接更新迭代点, 而是考虑从该点出发, 进一步减小二次模型. 也即, 在保证积极集 I_k^1 和 I_k^2 不变的情况下, 极小化二次函数

$$\psi_k(x_k^1 + d) = f_k + (g_k + B_k s_k^1)^T d + \frac{1}{2} d^T B_k d.$$

为了保证试探步的可行性, 我们借用传统的仿射变换方法的思想, 定义一个完全包含在可行域中的椭圆形信赖域, 在其中极小化 $\psi_k(x_k^1 + d)$. 在定义这个椭圆形信赖域时, 我们采用了下面这个仿射变换对角矩阵:

$$(\tilde{D}_k)_{ii} = \begin{cases} \min\{(x_k^1)^{(i)} - l^{(i)}, u^{(i)} - (x_k^1)^{(i)}, \Delta_k\}, & i \in I \setminus (I_k^1 \cup I_k^2), \\ 0, & i \in I_k^1 \cup I_k^2. \end{cases} \quad (2.9)$$

(2.9) 中引入的 Δ_k 与 (2.3) 有相同的作用. 现在, 我们定义带椭圆信赖域的 QP 子问题如下:

$$\min \quad (g_k + B_k s_k^1)^T d + \frac{1}{2} d^T B_k d \quad (2.10a)$$

$$\text{s.t.} \quad d^{(i)} = 0, \quad i \in I_k^1 \cup I_k^2, \quad (2.10b)$$

$$\sum_{i \notin I_k^1 \cup I_k^2} [(\tilde{D}_k)_{ii}^{-1} d^{(i)}]^2 \leq 1. \quad (2.10c)$$

记它的解为 s_k^2 . 由 $\tilde{D}_k$ 的定义 (2.9) 可知

$$\|s_k^2\| \leq \Delta_k. \quad (2.11)$$

定义

$$s_k = s_k^1 + s_k^2. \quad (2.12)$$

由 (2.5) 和 (2.11) 可知

$$\|s_k\| \leq 2\Delta_k.$$

基于 f 得到的真实下降量和基于二次模型得到的预测下降量之间的比值 (下降比) 为

$$\rho_k = \frac{f(x_k) - f(x_k + s_k)}{\psi_k(x_k) - \psi_k(x_k + s_k)}. \quad (2.13)$$

给定常数 $\eta \in (0, 1)$, 更新迭代点如下,

$$x_{k+1} = \begin{cases} x_k + s_k, & \rho_k \geq \eta, \\ x_k, & \rho_k < \eta. \end{cases} \quad (2.14)$$

在给出总的算法框架之前, 定义投影梯度

$$\hat{g}(x) = x - P[x - g(x), l, u],$$

其中投影算子 $P[x, l, u]$ 定义为

$$P[x, l, u]^{(i)} = \begin{cases} l^{(i)}, & x^{(i)} < l^{(i)}, \\ x^{(i)}, & l^{(i)} \leq x^{(i)} \leq u^{(i)}, \\ u^{(i)}, & x^{(i)} > u^{(i)}. \end{cases}$$

则 (2.1) 成立等价于

$$\hat{g}(x_*) = 0.$$

在下面的内容中, 采用 $\|\hat{g}(x)\|$ 作为算法终止与否的判断准则 (见文献 [21, 2]).

下面给出解一般约束优化问题的算法.

算法 2.1

步骤 1 初始化

给定可行的初始点 x_0 及初始信赖域半径 $\Delta_0 > 0$; 给定常数 $\varepsilon, \eta, \eta_1$ 和 η_2 满足

$$\varepsilon > 0, \quad 0 < \eta < \eta_1 \leq \eta_2 < 1.$$

计算 $f(x_0)$ 和 $\hat{g}(x_0)$, 并置 $k := 0$.

步骤 2 终止准则判断

若 $\|\hat{g}_k\| < \varepsilon$, 算法终止, 返回 x_k .

步骤 3 计算试探步

通过 (2.12) 计算试探步 s_k .

步骤 4 迭代点更新

计算 $f(x_k + s_k)$, 并通过 (2.13) 计算 ρ_k . 通过 (2.14) 计算 x_{k+1} .

步骤 5 信赖域半径更新

更新信赖域半径

$$\Delta_{k+1} = \begin{cases} \min \left\{ \frac{1}{2} \Delta_k, \max \{ \|s_k^1\|, \|s_k^2\| \} \right\}, & \rho_k < \eta_1, \\ \Delta_k, & \eta_1 \leq \rho_k \leq \eta_2, \\ \max \{ \Delta_k, 4 \max \{ \|s_k^1\|, \|s_k^2\| \} \}, & \rho_k > \eta_2. \end{cases}$$

第 4 节将把该算法应用到数值实验中并展开详细讨论.

3 全局收敛性分析

本节将证明算法 2.1 产生的迭代点列 $\{x_k\}$ 是全局收敛的, 并且每个聚点都是一阶稳定点. 在给出收敛性结果之前, 先给出几个假设.

假设 1 $f \in \mathbb{C}^2$.

假设 2 给定可行的初始点 x_0 , 水平集 $\mathcal{L} = \{x \in \mathcal{F} : f(x) \leq f(x_0)\}$ 是紧的.

假设 3 存在正数 M 和 G 使得, 对所有的 k 及所有的 $x \in \mathcal{L}$, $\|B_k\| \leq M$, $\|\nabla^2 f(x)\| \leq M$, $\Delta_k \leq G$ 及 $\|g_k\| \leq G$ 均成立.

由算法 2.1 的构建可以很自然地得到下面的引理.

引理 1 假设点列 $\{x_k\}$ 由算法 2.1 产生, 则

$$\{x_k\} \subseteq \mathcal{F}. \quad (3.1)$$

证明 对每个迭代点 $x_k \in \mathcal{F}$, 由 s_k^1 和 s_k^2 的定义 (2.5) 和 (2.10), 以及 $\tilde{D}_k$ 的定义可知

$$x_k + s_k \in \mathcal{F}.$$

归纳可得 $\{x_k\} \subseteq \mathcal{F}$. □

接下来我们来研究试探步 s_k 的性质. 先考虑 s_k^1 . 首先引入一个新的向量 w_k , 定义为

$$w_k^{(i)} = \min\{v_k^{(i)}, |g_k^{(i)}|\}. \quad (3.2)$$

由假设 3 可知

$$w_k^{(i)} \leq G. \quad (3.3)$$

引理 2 在假设 1-3 下, 沿 s_k^1 得到的预测下降量满足

$$-q_k(s_k^1) \geq \frac{1}{2} \frac{\|w_k \circ g_k\|}{\sqrt{n}} \min \left\{ \frac{\|w_k \circ g_k\|}{\sqrt{n}MG^2}, 1, \frac{\Delta_k}{G} \right\}, \quad (3.4)$$

这里

$$q_k(d) = g_k^T d + \frac{1}{2} d^T B_k d, \quad (3.5)$$

且“ $\circ$ ”表示向量之间的 Hadamard 乘积.

证明 首先, 由 s_k^1 的定义可知

$$-q_k(s_k^1) \geq -q_k \left(-\tau \frac{D_k^2 g_k}{\|D_k g_k\|} \right), \quad \forall \tau \in [0, 1].$$

由 D_k 的定义 (2.3) 知

$$\tau \left\| \frac{D_k^2 g_k}{\|D_k g_k\|} \right\| \leq \|D_k\| \leq \Delta_k, \quad x_k - \tau \frac{D_k^2 g_k}{\|D_k g_k\|} \in \mathcal{F}, \quad \forall \tau \in [0, 1].$$

定义函数 $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ 为

$$\varphi(\tau) = q_k \left(-\tau \frac{D_k^2 g_k}{\|D_k g_k\|} \right).$$

记 τ_* 是 φ 在 $[0, 1]$ 上的最小点, 且由 $\varphi(\tau)$ 的定义知

$$\varphi(\tau) = -\tau \|D_k g_k\| + \frac{\tau^2}{2} \frac{g_k^T D_k^2 B_k D_k^2 g_k}{\|D_k g_k\|^2}.$$

若 $\tau_* \in (0, 1)$, 则

$$\varphi(\tau_*) = -\frac{1}{2} \frac{\|D_k g_k\|^4}{g_k^T D_k^2 B_k D_k^2 g_k} \leq -\frac{1}{2} \frac{\|D_k g_k\|^2}{\|D_k B_k D_k\|}.$$

若 $\tau_* = 1$, 则由 $\varphi'(1) \leq 0$ 知

$$\varphi(\tau_*) \leq -\frac{\|D_k g_k\|}{2}.$$

由上述两个不等式知

$$-q_k(s_k^1) \geq \frac{1}{2} \min \left\{ \|D_k g_k\|, \frac{\|D_k g_k\|^2}{\|D_k B_k D_k\|} \right\}. \quad (3.6)$$

由

$$\|D_k g_k\| \geq \begin{cases} \min\{u^{(j)} - x_k^{(j)}, \Delta_k\} |g_k^{(j)}|, & \text{若 } g_k^{(j)} \leq 0, \\ \min\{x_k^{(j)} - l^{(j)}, \Delta_k\} |g_k^{(j)}|, & \text{若 } g_k^{(j)} > 0, \end{cases} \quad (3.7)$$

$$\|w_k \circ g_k\|_\infty = |w_k^{(i)} g_k^{(i)}|, \quad \text{对某个 } i,$$

及

$$\|D_k\| \leq \Delta_k \leq G$$

易知, 若 $g_k^{(i)} \leq 0$, 则

$$\begin{aligned} -q_k(s_k^1) &\geq \frac{1}{2} \min \left\{ \|D_k g_k\|, \frac{\|D_k g_k\|^2}{\|D_k B_k D_k\|} \right\} \\ &\geq \frac{1}{2} \min \left\{ (u^{(i)} - x_k^{(i)}) |g_k^{(i)}|, \frac{(u^{(i)} - x_k^{(i)})^2 |g_k^{(i)}|^2}{\|D_k B_k D_k\|}, \Delta_k |g_k^{(i)}|, \frac{\Delta_k^2 |g_k^{(i)}|^2}{\|D_k B_k D_k\|} \right\} \\ &\geq \frac{1}{2} \min \left\{ \|w_k \circ g_k\|_\infty, \frac{\|w_k \circ g_k\|_\infty^2}{MG^2}, \Delta_k \frac{\|w_k \circ g_k\|_\infty}{G}, \Delta_k^2 \frac{\|w_k \circ g_k\|_\infty^2}{MG^2 \Delta_k^2} \right\} \\ &= \frac{\|w_k \circ g_k\|_\infty}{2} \min \left\{ \frac{\|w_k \circ g_k\|_\infty}{MG^2}, 1, \frac{\Delta_k}{G} \right\} \\ &\geq \frac{\|w_k \circ g_k\|}{2\sqrt{n}} \min \left\{ \frac{\|w_k \circ g_k\|}{\sqrt{n}MG^2}, 1, \frac{\Delta_k}{G} \right\}. \end{aligned}$$

对于 $g_k^{(i)} > 0$ 可得相同的结果. □

由上述引理可得引理 3.

引理 3 在假设 1-3 下, 若 $\{s_k\}$ 满足

$$q_k(s_k) \leq q_k(s_k^1), \quad \|s_k\| \leq 2\Delta_k, \quad \text{且 } x_k + s_k \in \mathcal{F},$$

则

$$\liminf_{k \rightarrow \infty} \|w_k \circ g_k\| = 0. \quad (3.8)$$

证明 采用反证法. 假设存在 k_0 及 $\varepsilon > 0$ 使得

$$\|w_k \circ g_k\| \geq \varepsilon, \quad \forall k \geq k_0.$$

记 $\mathcal{S}$ 是所有成功迭代的指标, 即

$$\mathcal{S} = \{k \in \mathbb{Z}_+ : \rho_k \geq \eta\}.$$

由引理 2 知, 对任意 $k \geq k_0, k \in \mathcal{S}$,

$$\begin{aligned} -q_k(s_k) &\geq -q_k(s_k^1) \geq \frac{1}{2} \frac{\|w_k \circ g_k\|}{\sqrt{n}} \min \left\{ \frac{\|w_k \circ g_k\|}{\sqrt{n}MG^2}, 1, \frac{\Delta_k}{G} \right\} \\ &\geq \frac{1}{2} \frac{\varepsilon}{\sqrt{n}} \min \left\{ \frac{\varepsilon}{\sqrt{n}MG^2}, 1, \frac{\Delta_k}{G} \right\}. \end{aligned}$$

若 $\mathcal{S}$ 是一个有限集合, 由算法 2.1 的构建过程可知对所有充分大的 k , 成立

$$\Delta_{k+1} \leq \frac{1}{2} \Delta_k,$$

从而

$$\sum_{k=1}^{\infty} \Delta_k < \infty.$$

若 $\mathcal{S}$ 是一个无限集合, 由 (2.14) 知, 对所有 $k \in \mathcal{S}$ 都有

$$f_k - f_{k+1} \geq -\eta q_k(s_k),$$

又 $\{f_k\}$ 单调下降, 则

$$\sum_{k \in \mathcal{S}} \Delta_k < \infty.$$

由信赖域半径的更新策略可知

$$\sum_{k=1}^{\infty} \Delta_k \leq \left(1 + \frac{4}{1 - \frac{1}{2}}\right) \left(\Delta_1 + \sum_{k \in \mathcal{S}} \Delta_k\right) < \infty.$$

所以

$$\lim_{k \rightarrow \infty} \Delta_k = 0. \quad (3.9)$$

进而, 对所有充分大的 $k \in \mathcal{S}$ 有

$$-q_k(s_k) \geq \frac{\varepsilon}{2\sqrt{n}G} \Delta_k, \quad (3.10)$$

以及

$$\lim_{k \rightarrow \infty} \|x_k - x_{k+1}\| = 0.$$

由假设易知, 对所有充分大的 k ,

$$|f_k - f_{k+1} + q_k(s_k)| \leq \frac{1}{2} |s_k^T (\nabla^2 f(\xi_k) - B_k) s_k| \leq M \|s_k\|^2 \leq 4M \Delta_k^2, \quad (3.11)$$

其中 $\xi_k \in [x_k, x_{k+1}]$. 结合 (3.10) 以及 (3.11), 可得下降比 ρ_k 和 1 之间的关系:

$$|\rho_k - 1| = \left| \frac{f_k - f_{k+1}}{-q_k(s_k)} - 1 \right| \leq \frac{8\sqrt{n}MG}{\varepsilon} \Delta_k.$$

再由 (3.9) 知

$$\lim_{k \rightarrow \infty} \rho_k = 1,$$

故而对所有充分大的 k , 有

$$\Delta_{k+1} \geq \Delta_k,$$

这与 (3.9) 矛盾. 所以 (3.8) 成立. $\square$

上述引理证明序列 $\{w_k \circ g_k\}$ 的下极限是 0, 实际上, 我们可以得到更强的收敛性结果. 基于上面的分析, 下面给出本文最主要的收敛性结果.

定理 1 在假设 1-3 下, 若 $\{s_k\}$ 满足

$$q_k(s_k) \leq q_k(s_k^1), \quad \|s_k\| \leq 2\Delta_k, \quad \text{且 } x_k + s_k \in \mathcal{F},$$

则

$$\lim_{k \rightarrow \infty} \|w_k \circ g_k\| = 0. \quad (3.12)$$

所以, $\{x_k\}$ 的每个聚点都是一阶稳定点.

证明 采用反证法. 假设存在子列 $\{m_i\} \subset \mathcal{S}$ 及 $\varepsilon > 0$, 使得

$$\|w_{m_i} \circ g_{m_i}\| \geq 2\varepsilon. \quad (3.13)$$

由引理 3.3 知

$$\liminf_{k \rightarrow \infty} \|w_k \circ g_k\| = 0,$$

这保证了 m_i 的存在性, 记比 m_i 大的, 同时满足 $\|g_{n(m_i)}\| < \varepsilon$ 的, 且为成功迭代的指标为 $n(m_i)$. 简单起见, 记 $n_i \triangleq n(m_i)$, 由此我们得到了子列 $\{n_i\} \subseteq \mathcal{S}$ 满足

$$\|w_k \circ g_k\| \geq \varepsilon \quad (m_i \leq k < n_i) \quad \text{及} \quad \|w_{n_i} \circ g_{n_i}\| < \varepsilon. \quad (3.14)$$

考虑集合

$$\mathcal{K} \triangleq \{k \in \mathcal{S} : m_i \leq k < n_i\}.$$

由充分下降性条件知, 对所有的 $k \in \mathcal{K}$,

$$\begin{aligned} f_k - f_{k+1} &\geq -\eta q_k(s_k) \geq \frac{1}{2}\eta \frac{\|w_k \circ g_k\|}{\sqrt{n}} \min \left\{ \frac{\|w_k \circ g_k\|}{\sqrt{n}MG^2}, 1, \frac{\Delta_k}{G} \right\} \\ &\geq \frac{1}{2}\eta \frac{\varepsilon}{\sqrt{n}} \min \left\{ \frac{\varepsilon}{\sqrt{n}MG^2}, 1, \frac{\Delta_k}{G} \right\}. \end{aligned}$$

由于 f 有下界, 则

$$\lim_{\substack{k \rightarrow \infty \\ k \in \mathcal{K}}} \Delta_k = 0.$$

于是对所有充分大的 $k \in \mathcal{K}$,

$$f_k - f_{k+1} \geq \frac{\eta\varepsilon}{2\sqrt{n}G} \Delta_k \triangleq \bar{\varepsilon} \Delta_k. \quad (3.15)$$

利用 (3.15) 及三角不等式知, 对所有充分大的 i ,

$$\|x_{m_i} - x_{n_i}\| \leq \sum_{k=m_i}^{n_i-1} \|x_k - x_{k+1}\| \leq 2 \sum_{k=m_i}^{n_i-1} \Delta_k \leq \frac{2}{\bar{\varepsilon}} [f(x_{m_i}) - f(x_{n_i})].$$

由序列 $\{f_k\}$ 的单调下降性知, 上式右端趋于 0, 所以

$$\lim_{i \rightarrow \infty} \|x_{m_i} - x_{n_i}\| = 0.$$

再结合假设 1 知

$$\lim_{i \rightarrow \infty} \|g_{m_i} - g_{n_i}\| = 0, \quad \lim_{i \rightarrow \infty} \|W_{m_i} - W_{n_i}\| = 0,$$

这里 $W = \text{diag}(w)$. 所以存在 i_0 使得对所有的 $i \geq i_0$,

$$\|g_{m_i} - g_{n_i}\| < \frac{\varepsilon}{2G}, \quad \|W_{m_i} - W_{n_i}\| < \frac{\varepsilon}{2G}.$$

结合 (3.14) 可知, 对所有的 $i \geq i_0$ 都有

$$\begin{aligned} \|w_{m_i} \circ g_{m_i}\| &\leq \|W_{m_i}\| \|g_{m_i} - g_{n_i}\| + \|(W_{m_i} - W_{n_i})g_{n_i}\| + \|w_{n_i} \circ g_{n_i}\| \\ &\leq G \|g_{m_i} - g_{n_i}\| + \|(W_{m_i} - W_{n_i})g_{n_i}\| + \|w_{n_i} \circ g_{n_i}\| \\ &< 2\varepsilon, \end{aligned}$$

这与 (3.13) 矛盾. 故而

$$\lim_{k \rightarrow \infty} \|w_k \circ g_k\| = 0.$$

因此 (3.12) 成立.

假设 x_* 是点列 $\{x_k\}$ 的任一聚点, 则

$$\min\{v^{(i)}(x_*), |g^{(i)}(x_*)|\} |g^{(i)}(x_*)| = 0, \quad i = 1, \dots, n. \quad (3.16)$$

由此可知, 对任意 i , 若 $l^{(i)} < x_*^{(i)} < u^{(i)}$, 一定有

$$g_*^{(i)} = 0.$$

若 $x_*^{(i)} = l^{(i)}$, 则

$$g_*^{(i)} \geq 0.$$

否则, 如果 $g_*^{(i)} < 0$, 则对所有充分大的 k , 均有 $g_k^{(i)} < 0$, 于是从 (2.2) 中 $v(x)$ 的定义可知,

$$v_k^{(i)} = u^{(i)} - x_k^{(i)} \geq \frac{u^{(i)} - l^{(i)}}{2},$$

这与 (3.16) 矛盾. 类似地, 若 $x_*^{(i)} = u^{(i)}$, 则有 $g_*^{(i)} \leq 0$. □

4 数值试验

本节考察算法 2.1 的数值表现. 程序由 Matlab 7.6.0 (R2008a) 编写. 本节数值结果的测试环境是 PC 机 (1.86 GHz, 1GB of RAM) Pentium 双核处理器, Fedora 8.0 操作系统. 测试了 CUTEr^[22] 里面的所有 106 个带有简单界约束的优化问题. 由于电脑内存的限制, 只有 103 个问题被成功解码.

CUTEr 能够提供函数的精确 Hesse 阵, 所以在子问题 (2.10) 中令 $B_k = \nabla^2 f(x_k)$, 并重写 (2.10) 为

$$\min \quad \bar{q}_k(d) = [\tilde{D}_k(g_k + B_k d_k)]^T \tilde{d} + \frac{1}{2} \tilde{d}^T \tilde{D}_k B_k \tilde{D}_k \tilde{d} \quad (4.1a)$$

$$\text{s.t.} \quad \|\tilde{d}\| \leq 1. \quad (4.1b)$$

调用 Moré 的子程序 gqptar^[23] 得到 (4.1) 的解 $\tilde{d}_k$. 有了 $\tilde{d}_k$ 后, 并不直接令 $s_k^2 = \tilde{D}_k \tilde{d}_k$, 而是在信赖域约束之下, 沿着 $\tilde{D}_k \tilde{d}_k$ 尽可能地减小二次模型, 置 s_k^2 为

$$s_k^2 = \arg \min_d \{\tilde{q}_k(d) : d = \tau \tilde{D}_k \tilde{d}_k, \|d\| \leq \Delta_k, x_k + d \in \mathcal{F}\}.$$

这样定义的 s_k^2 依然可以保证算法的收敛性.

其次, CUTEr 能够提供初始点 x_{start} , 但是该点不一定是可行的, 所以我们选取可行的初始点如下:

$$x_0^{(i)} = \begin{cases} l^{(i)}, & \text{若 } x_{\text{start}}^{(i)} < l^{(i)}, \\ u^{(i)}, & \text{若 } x_{\text{start}}^{(i)} > u^{(i)}, \\ x_{\text{start}}^{(i)}, & \text{否则.} \end{cases}$$

再有, CUTEr 的问题中, 有些变量取值是固定的, 也即上下界是相等的. 针对这种情形, 令 $\mathcal{I}$ 是所有那些固定变量的指标集合. 令 e_i 是 n 阶单位阵的第 i 列. 令 I 是由 e_i ($i \notin \mathcal{I}$) 排列起来得到的矩阵. 则将 (2.10) 视为自由变量 $\bar{d} = I^T d$ 的函数.

算法中用到的参数定义如下:

$$\Delta_0 = 1, \quad \eta = 10^{-8}, \quad \eta_1 = 0.2, \quad \eta_2 = 0.8, \quad \varepsilon = 10^{-5}.$$

由于 (2.7) 在每步迭代中确定的积极集数可能很少, 从而可能导致收敛速度变慢, 所以在实际计算中, 采用如下的积极集定义方式:

$$I_k^1 = \{i : (x_k^1)^{(i)} - l^{(i)} \leq 10^{-4} \Delta_k, g^{(i)}(x_k^1) \geq 0\},$$

$$I_k^2 = \{i : u^{(i)} - (x_k^1)^{(i)} \leq 10^{-4} \Delta_k, g^{(i)}(x_k^1) \leq 0\}.$$

由于梯度的计算代价可能会很大, 所以并不直接计算 $g(x_k^1)$, 而是通过计算 $g_k + B_k s_k^1$ 来近似. 在算法 2.1 中做上述修改之后, 得到算法 2.1 mod, 并在表 1 中, 列出了这个新算法和 L-BFGS-B 的数值结果. 对每个问题, 记 f 表示目标函数值的计算次数, g 表示梯度值的计算次数. 同时, n 表示问题规模. 记“F”表示算法非正常终止, 即迭代 1000 步以后仍未找到解.

我们与求解界约束优化问题的著名软件 L-BFGS-B^[24] 进行了数值比较. 公平起见, 要求两个算法涉及的所有参数取值相同. 并且若 L-BFGS-B 在迭代 1000 次后仍没给出解, 也认为它非正常终止. 为了清晰地表达算法之间的比较, 我们绘制了算法的数值表现图, 这种思想最早在文献 [25] 中提出.

从图 1 可看出, 算法 2.1 mod 要比算法 2.1 的数值表现要好一些. 算法 2.1 mod 对约 85% 的问题所需要的函数值计算次数要少, 对约 90% 的问题所需要的梯度值计算次数要少. 同时, 有 90% 的问题在执行算法 2.1 mod 时, 所用的函数值次数比最少值的 2 倍要少. 图 2 给出了算法 2.1 mod 与 LBFGS-B 的比较, 类似于前面的分析可知, 算法 2.1 对于测试问题在函数值计算次数以及梯度值计算次数方面都要比 L-BFGS-B 优越.

5 结论

本文提出了解一般界约束优化问题的一种新的积极集信赖域方法. 为了保证全局收敛性, 我们首先沿着一个尺度变换后的梯度方向进行搜索, 然后记录下相应的积极集. 这里定义的仿射变换矩阵不仅依赖于变量到边界的距离, 还依赖于当前点的梯度值以及该步迭代的信赖域半径. 为了进一步加快收敛速度, 在积极集的零空间里面进一步减小二次模型, 从而得到新的试探点, 最后基于该点更新迭代点. 通过分析可知, 算法产生的迭代点列是全局收敛的, 并且每个聚点都是一阶稳定点. 最后我们给出了算法的数值表现. 通过测试 CUTEr 算例库中所有的界约束优化问题, 并与 L-BFGS-B 进行了比较. 数值结果进一步验证了前面的理论分析, 并且表明我们提出的算法还是非常有效的.

表 1 数值结果

问题	n	新方法		LBFGSB		问题	n	新方法		LBFGSB	
		f	g	f	g			f	g	f	g
3PK	100	34	34	34	34	OSLBQP	8	4	4	3	3
ALLINIT	4	7	6	16	16	PALMER1	4	25	13	6	6
BDEXP	1000	15	15	18	18	PALMER1A	6	49	34	7	7
BIGGSB1	100	95	95	332	332	PALMER1B	4	24	20	67	67
BLEACHNG	9	F	F	F	F	PALMER1E	8	84	56	816	816
BQP1VAR	1	2	2	2	2	PALMER2	4	18	11	6	6
BQPGABIM	46	7	5	4	4	PALMER2A	6	93	67	138	138
BQPGASIM	50	8	6	4	4	PALMER2B	4	33	20	38	38
CAMEL6	2	6	6	15	15	PALMER2E	8	103	55	630	630
CHEBYQAD	100	108	62	3	3	PALMER3	4	23	11	221	221
CVXBQP1	1000	6	6	2	2	PALMER3A	6	105	65	155	155
DECONVB	61	34	23	29	29	PALMER3B	4	15	12	49	49
EG1	3	6	6	9	9	PALMER3E	8	103	58	570	570
EXPLIN	1200	55	55	212	212	PALMER4	4	17	16	48	48
EXPLIN2	1200	161	111	93	93	PALMER4A	6	64	43	123	123
EXPQUAD	120	35	35	17	17	PALMER4B	4	15	11	44	44
GRIDGENA	1226	5	5	21	21	PALMER4E	8	75	45	582	582
HADAMALS	400	189	184	20	20	PALMER5A	8	F	F	F	F
HART6	6	8	6	17	17	PALMER5B	9	944	477	F	F
HATFLDA	4	27	27	12	12	PALMER5D	8	9	9	30	30
HATFLDB	4	24	24	68	68	PALMER5E	8	F	F	F	F
HATFLDC	25	5	5	23	23	PALMER6A	6	144	93	251	251
HIMMELP1	2	8	8	12	12	PALMER6E	8	84	49	458	458
HS1	2	32	23	47	47	PALMER7A	6	F	F	F	F
HS110	10	6	5	7	7	PALMER7E	8	920	471	F	F
HS2	2	10	8	15	15	PALMER8A	6	38	26	110	110
HS25	3	1	1	1	1	PALMER8E	8	49	30	338	338
HS3	2	4	4	4	4	PENTDI	1000	2	2	3	3
HS38	4	57	38	33	33	PROBPENL	500	4	4	3	3
HS3MOD	2	5	5	9	9	PSPDOC	4	7	6	11	11
HS4	2	2	2	2	2	QR3DLS	155	64	54	F	F
HS45	5	4	4	11	11	S368	100	58	6	21	21
HS5	2	5	5	8	8	SCOND1LS	502	556	517	47	47
JNLBRNG1	1024	27	27	5	5	SIM2BQP	1	2	2	2	2
JNLBRNG2	1024	13	13	5	5	SIMBQP	2	5	5	6	6
JNLBRNGA	1024	19	19	81	81	SINEALI	1000	14	12	13	10
JNLBRNGB	1024	46	46	485	485	SPECAN	9	9	9	85	85
LINVERSE	199	38	26	36	36	TORSION1	1024	29	29	40	40

表 1 (续) 数值结果

问题	n	新方法		LBFGSB		问题	n	新方法		LBFGSB	
		f	g	f	g			f	g	f	g
LOGROS	2	80	50	118	118	TORSION2	1024	30	30	24	24
MAXLIKA	8	14	12	21	21	TORSION3	1024	11	11	23	23
MCCORMCK	1000	8	8	15	15	TORSION4	1024	32	32	40	40
MDHOLE	2	59	37	89	89	TORSION5	1024	9	9	14	14
NCVXBQP1	1000	22	22	2	2	TORSION6	1024	32	31	35	35
NCVXBQP2	1000	39	39	61	61	TORSIONA	1024	29	29	40	40
NCVXBQP3	1000	38	38	83	83	TORSIONB	1024	27	26	94	94
NOBNDTOR	1024	29	28	14	14	TORSIONC	1024	12	12	21	21
NONSCOMP	500	46	34	29	29	TORSIOND	1024	36	36	72	72
OBSTCLAE	1024	40	39	63	63	TORSIONE	1024	9	9	13	13
OBSTCLAL	1024	14	14	48	48	TORSIONF	1024	29	28	38	38
OBSTCLBL	1024	110	110	63	63	WEEDS	3	35	22	80	80
OBSTCLBM	1024	36	36	39	39	YFIT	3	48	38	99	99
OBSTCLBU	1024	36	34	43	43						

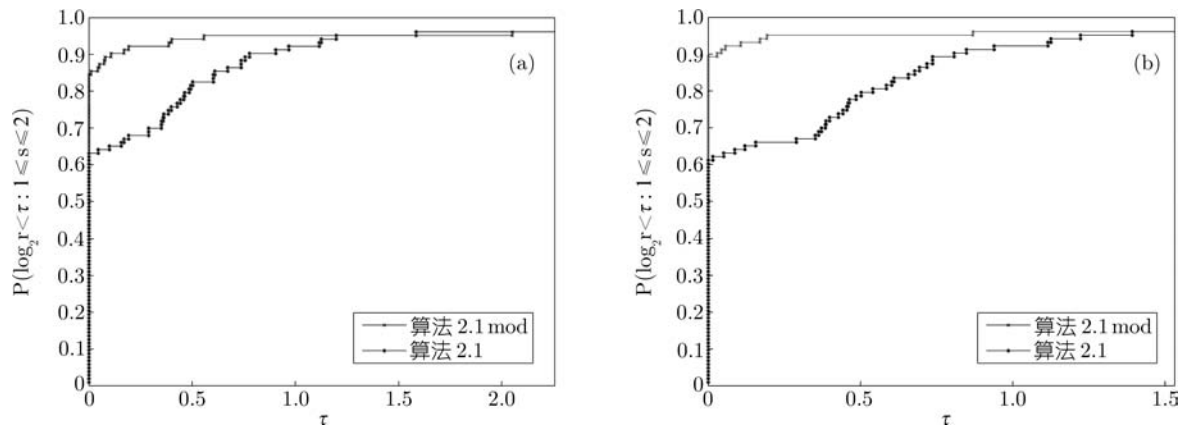


图 1 算法 2.1 mod 和算法 2.1 的函数值计算次数 (a) 和梯度值计算次数 (b) 的性能表现

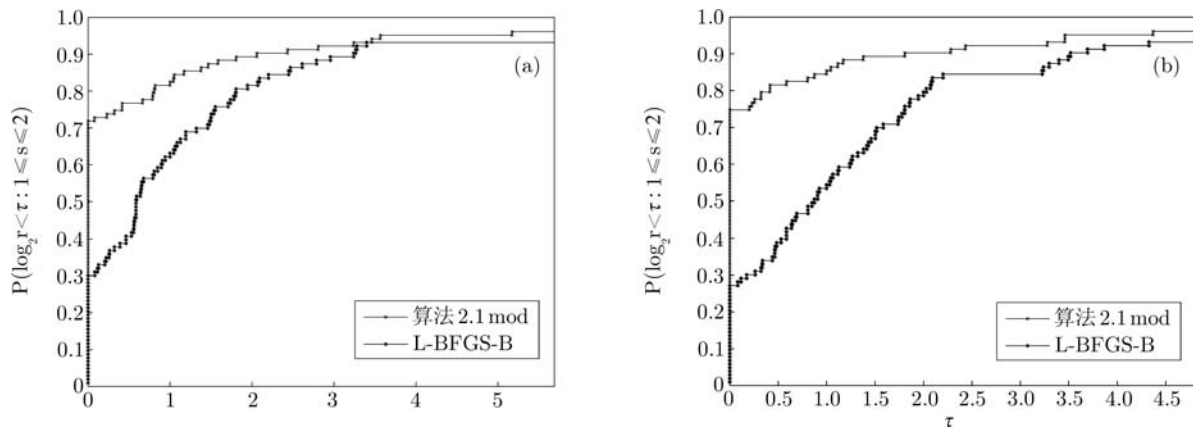


图 2 算法 2.1 mod 和 L-BFGS-B 的函数值计算次数 (a) 和梯度值计算次数 (b) 的性能表现

致谢 作者诚挚感谢其导师袁亚湘研究员对本文提出的诸多宝贵的修改意见. 同时对审稿人提出的重要意见表示衷心的感谢.

参考文献

- 1 Hager W W, Zhang H. A new active set algorithm for box constrained optimization. *SIAM J Optim*, 2006, 17: 526–557
- 2 Conn A R, Gould N I M, Toint Ph L. Global convergence of a class of trust region algorithms for optimization with simple bounds. *SIAM J Numer Anal*, 1988, 25: 433–460; 1989, 26: 764–767
- 3 Lin C, Moré J J. Newton's method for large bound-constrained optimization problem. *SIAM J Optim*, 1999, 9: 1100–1127
- 4 Dikin I I. Iterative solution of problems of linear and quadratic programming. *Doklady Akademii Nauk SSSR*, 1967, 174: 747–748. Translated in: *Soviet Math Doklady*, 1967, 8: 674–675
- 5 Barnes E R. A variation on Karmarkar's algorithm for solving linear programming problems. *Math Program*, 1986, 36: 174–182
- 6 Vanderbei R J, Meketon M S, Freedman B A. A modification of Karmarkar's linear programming algorithm. *Algorithmica*, 1986, 1: 395–407
- 7 Dikin I I, Zorkaltsev V I. *Iterative Solutions of Mathematical Programming Problems*. Novosibirsk: Nauka, 1980
- 8 Bonnans J, Bouhtou M. The trust region affine interior point algorithm for convex and nonconvex quadratic programming. *RAIRO Oper Res*, 1995, 29: 195–217
- 9 Monteiro R D C, Tsuchiya T. Global convergence of the affine scaling algorithm for convex quadratic programming. *SIAM J Optim*, 1998, 8: 26–58
- 10 Tseng P. Convergence properties of Dikin's affine scaling algorithm for nonconvex quadratic minimization. *J Global Optim*, 2004, 30: 285–300
- 11 Ye Y. On an affine scaling algorithm for nonconvex quadratic programming. *Math Program*, 1992, 56: 285–300
- 12 Monteiro R D C, Wang Y. Trust region affine scaling algorithms for linearly constrained convex and concave programs. *Math Program*, 1998, 80: 283–313
- 13 Sun J. A convergence analysis for a convex version of Dikin's algorithm. *Ann Oper Res*, 1996, 62: 357–374
- 14 Coleman T F, Li Y. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM J Optim*, 1996, 6: 418–445
- 15 Powell M J D. A method for nonlinear constraints in minimization problems. In: Fletcher R, ed. *Optimization*. London: Academic Press, 1969
- 16 Vardi A. A trust region algorithm for equality constrained minimization: convergence properties and implementation. *SIAM J Numer Anal*, 1985, 22: 575–591
- 17 Yuan Y. On the convergence of a new trust region algorithm. *Numer Math*, 1995, 70: 515–539
- 18 Nocedal J, Wright S J. *Numerical Optimization*. New York: Springer, 2006
- 19 Moré J J. Recent developments in algorithms and software for trust region methods. In: Bachem M G A, Korte B, eds. *The State of the Art. Math Program*. Berlin: Springer-Verlag, 1983, 259–287
- 20 Yuan Y. *Computational Methods for Nonlinear Optimization (in Chinese)*. Beijing: Science Press, 2007
- 21 Conn A R, Gould N I M, Toint Ph L. *Trust-region Methods*. MPS-SIAM Series on Optimization, vol. 1. Philadelphia: SIAM, 2000
- 22 Gould N I M, Orban D, Toint Ph L. CUTer, a constrained and unconstrained testing environment, revisited. *ACM Trans Math Software*, 2003, 29: 373–394
- 23 Moré J J, Sorensen D C. Computing a trust region step. *SIAM J Sci Comput*, 1983, 4: 553–572
- 24 Zhu C, Byrd R H, Nocedal J. L-BFGS-B, FORTRAN routines for large scale bound constrained optimization. *ACM Trans Math Software*, 1997, 23: 550–560
- 25 Dolan E D, Moré J J. Benchmarking optimization software with performance profiles. *Math Prog*, 2002, 91: 201–213

An active set trust region method for general bound constrained optimization

WANG Xiao

Abstract We study a new active set trust region method for general bound constrained optimization problems. Each iteration is divided into two stages. At the first stage, from x_k we take a scaled gradient step to obtain a trial point x_k^1 at which we record its active constraints. Here the affine scaling matrix is dependent on the distances between x_k and its boundaries, its gradient and also the trust region radius. At the second stage, a trial step is taken from x_k^1 by solving a QP subproblem with an ellipsoid constraint defined in the null space of all the active constraints. Then we get the new trial point and judge whether or not to set it as the next iterate. We obtain that every accumulation point of the iterates satisfies the first-order optimality condition. Numerical experiment for problems with bound constraints from the famous CUTer collection is also reported. The numerical performance reveals that our method is effective and comparable with L-BFGS-B.

Keywords: bound constrained optimization, active set, affine scaling, trust region

MSC(2000): 65K03

doi: 10.1360/012010-266