第28卷第4期　　上海师范大学学报(自然科学版)　　Vol 28, No. 4
1999年12月　　J. of Shanghai Teachers Univ. (Natural Sciences)　　Dec. 1999

A Trust Region Algorithm for Nonsmooth Unary Optimization

ZHU De-tong

(College of Mathematical Science, Shanghai Teachers University, Shanghai, 200234, China)

Abstract: Presents a trust region algorithm for nonsmooth unary optimization problems. Based on the duality theorem of linear programming, the directional derivatives of the objective function can be expressed as a linear programming which is very important in the practical calculution of trust region subproblems. Gives a theoretical analysis which proves that the proposed algorithm is globally convergent and has a local superlinear rate under some reasonable conditions.

Key words: Trust region; unary optimization; convergence

CLC number: O221.2　　**Document code**: A　　**Article ID**: 1000-5137(1999)04-0001-09

1 Introduction

In this paper we consider the nonlinear unconstrained minimization problems

$$\min f(x), \tag{1.1}$$

where $f(x): R^n \to R^1$ is convex function. Following Goldfurb and Wang in [2] and MoCormick and Sofer [3] we call problem (1.1) a unary convex optimization problem if $f(x)$ takes the form

$$f(x) = \sum_{i=1}^{m} U_i(\alpha_i(x)), \tag{1.2}$$

where for $i = 1, \ldots, m$, $m \geq n$, $\alpha_i(x) = a_i^T x$, a_i is a constant vectors of size $n \times 1$ and $U_i(\ )$ is a unary convex function, i. e., $U_i(\alpha_i): R^1 \to R^1$, not necessarily differentiable (e. g., piecewise linear or quadratic). Note that the unary functions such as, the separable function, the objective function of the linear robust regrassion problem, the dual objective function of the

Received date: 1999-03-04

Foundation item: The Science and Technology Foundation of Shanghai Higher Education

Biography: ZHU De-tong (1954-), male, Associate Professor, doctor, College of Mathematical Science, Shanghai Teachers University.

entropy problem in information theory[3] and the dual objection function of the lower level programming in pipe network optimization[7] are of the unary convex function form. Many proposed algorithms for the unary optimization utilized the derivatives of unary functions involving twice continuously differentiable derivatives. However it is not always to be supposed, or even given as a finite function on all of R^1. Such as piecewise linear and piecewise quadratic. In such convex unary functions, there are special ways of generating directions of descent, and the duality of the linear programming can play a very strong role.

Trust region methods are an important class of iterative methods for solving nonlinear optimization problems to assure global convergence. In an unconstrained minimization problem, a step to a new iterate is obtained by minimizing a local model of the objective function over a restricted region centered about the current iterate. The size of this restricted region depends on how well the local model predicts the behavior of the objective function. This strategy will force global convergence of the iterates from an arbitrary starting point to a point which satisfies the first-order necessary conditions.

The main purpose of this paper is to propose trust region method with the generating directions of descent by adopting the duality idea. The paper is organized as follows. In section 2, we present the optimality and elementary directions of descent. In section 3, we describe the algorithm which combines the techniques of trust region and descent directions. In section 4, the global convergences of the proposed algorithm are established. The further superlinear convergence rate is discussed under some mild conditions in section 5.

2 Optimality and Elementary Directions of Descent

Just as linear programming problem can be descried qualitatively as one where a linear function is minimized over a convex polyhedron, so can a unary convex programming problem be described as one of the form

$$\text{(P)} \qquad \min_{x \in \Omega} f(x) = \sum_{i=1}^{m} U_i(\alpha_i(x)),$$

where Ω is a convex polyhedron in R^n, and for $i = 1, 2, \cdots, m$, $m \geq n$, $\alpha_i(x) = a_i^T x$, a_i is a constant vectors for size $n \times 1$ and $U_i(x)$ is a unary convex function.

To introduce the exact technical assumptions that will be needed, and to put problem (1.1) in a convenient "normalized" form at the same time, we specify now that for each $i = 1, \cdots, m$, we have

$$\text{a nonempty real interval } I_i, \tag{2.1}$$

not necessarily closed, possibly all of $R^1 = (-\infty, +\infty)$, and a convex function $U_i: I_i \to R^1$ which is continuous relative to I_i. (Regard U_i as $+\infty$ outside of I_i). Letting

$$\alpha_i = \alpha_i(x), \quad x = (x_1, \cdots, x_n)$$

observe from the linearingity of α_i that x ranges over a certain subspace $\bar{\Omega}$ of R^n. As x ranges over R^n as pointed by Rockafollar in [3], every unary programming problem can therefore be reduced to the more fundamental form

$$(\mathrm{P}) \qquad \min \quad f(x) = \sum_{i=1}^{m} U_i(\alpha_i)$$

$$\text{s.t.} \quad \alpha_v \in I_v \text{ for } v = 1, \cdots, m,$$

$$\alpha = (\alpha_1, \cdots, \alpha_n) \in \bar{\Omega},$$

where I_i and U_i are as in (2.1), (1.2), respectively for $i = 1, \cdots, m$ and $\bar{\Omega}$ is some subspace of R^n. This is a separable convex programming problem with linear constraints, some of which are given abstractly by the condition $x \in \Omega$, but can be represented in other ways as the situation warrants. For $i = 1, \cdots, m$, $\alpha_i(x) = a_i^T x$, we have that

$$\alpha = A x \tag{2.2}$$

where $\alpha = (\alpha_1, \cdots, \alpha_m)$ and $A^T = [a_1, \cdots, a_n]$. Equation (2.2) means that $\alpha \in R(A) = \{Ax \mid x \in R^n\}$, the range subspace of A. Since A has full rank m, we have $\dim(R(A)) = m$ and $\dim(N(A^T)) = n - m$ where $N(A^T)$ is the null subspace of A^T. Let $Q = [V, W] \in R^{n\times n}$ be an orthogomal matrix in which the columns of $V \in R^{n\times m}$ form an orthonormal basis of $R(A)$ and the columns $W \in R^{n\times(n-m)}$ form an orthonormal basis of $N(A^T)$. Clearly, by an elementary result in linear algebra, we have that $\alpha \in R(A)$ if and only if $W\alpha = 0$, i.e., the projection of the α onto $N(A^T)$ is zero. This establishes that the subspace $R(A)$ can be expressed as the set of all solutions of the linear system

$$\{\alpha \mid W\alpha = 0\} \tag{2.3}$$

Therefore, the unary programming problem can be reduced to the following fundamental form

$$(\mathrm{P}) \qquad \min \quad f(\alpha) = \sum_{i=1}^{m} U_i(\alpha_i)$$

$$\text{s.t.} \quad \{\alpha \mid W\alpha = 0\} \quad \alpha \in \bar{\Omega}$$

Let

$$\Omega = \bar{\Omega} \cap \{\alpha \mid W\alpha = 0\}, \tag{2.4}$$

since $\bar{\Omega}$ is a subspace on R^n, Ω is a convex polyhedrom in R^n. The problem can be reduced to a separable convex programming problem with linear constraints as follows

$$(\mathrm{P}) \qquad \min \quad f(\alpha) = \sum_{i=1}^{m} U_i(\alpha_i)$$

$$\text{s.t.} \quad \alpha \in \Omega$$

In problem (P), the minimal function $f(x)$ is a convex function defined on Ω. Thus x is a feasible solution to (P) if and only if it is a point of Ω where $f(x)$ is finite, and it is an optimal solution if and only if, in addition,

$$f^0(x; d) \geq 0 \text{ for all } d \in \Omega, \tag{2.5}$$

where of course, the generalized directional derivative of f at x in the direction $d \in R^n$,

$$f^0(x; d) = \lim_{y\to 0} \sup_{t\downarrow 0} \frac{f(y + td) - f(y)}{t}.$$

A function f is said to be regular at a if the one-sided directional derivative

$$f'(x;d) = \lim_{t\downarrow 0} \frac{f(y+td)-f(y)}{t}. \tag{2.6}$$

exists for all directions $d \in R^n$ and

$$f'(x;d) = f^0(x;d). \tag{2.7}$$

The function f is said to be regular on a set Ω if it is regular at every point of the set Ω. Since the unary function discussed in this paper is always assumed to be convex analysis (for example, see Theorem 23.1 on page 213 of Rockafellar [4]). A convex function always has a one-sided directional derivative. The convex analysis says that these two derivatives coincide for a convex function.

On the other hand for (2.1) case, any $d \in \Omega$ with $f'(x;d) < 0$, if one exists and $f^0(x;d) = f'(x;d)$, given a direction of descent from x; for small enough $t > 0$, $x + td$ is another feasible solution to (P) and $f(x+td) < f(x)$.

Each U_i has a right derivative $U'_{i+}(\alpha_i)$ and a left derivative $U'_{i-1}(\alpha_i)$ at every $\alpha_i \in R^1$. These are nondecreasing functions of α_i that satisfy

$$-\infty < U'_{i+}(\alpha_i) \le U'_{i-}(\alpha_i) < \infty. \tag{2.8}$$

As $f(x)$ is a unary convex function, its directional derivative along any direction d exists, which can be written by a direct formula for the directional derivative of $f(x)$,

$$f'(x;d) = \lim_{t\downarrow 0} \frac{f(y+td)-f(y)}{t} =$$

$$\lim_{t\downarrow 0} \Big(\sum_{i,d_i>0} \frac{U_i(\alpha_i+td_i)-U_i(\alpha_i)}{t} + \sum_{i,d_i<0} \frac{U_i(\alpha_i+td_i)-U_i(\alpha_i)}{t} \Big) =$$

$$\sum_{i,d_i>0} U'_{i+}(\alpha_i)d_i + \sum_{i,d_i<0} U'_{i-}(\alpha_i)d_i. \tag{2.9}$$

It is clear to see that the value of $f'(x;d)$ depends on sign of d_i so that its calculation is inconvenient.

For f itself, the generalized gradient of f at α is the set

$$\partial f(\alpha) = \{ g \in R^m \mid g^T d \le f^0(\alpha;d),\ \forall d \in R^m \} \tag{2.10}$$

and because of separability this reduces to

$$\partial f(\alpha) = \partial U_1(\alpha_1) \times \partial U_2(\alpha_2) \times \cdots \times \partial U_m(\alpha_m) \tag{2.11}$$

Furthermore, using the derivatives of $U_i(\alpha_i)$, $i = 1, 2, \cdots m$, we have that the subgradient of $U_i(\alpha_i)$,

$$\partial U_i(\alpha_i) = \{ t_i \in R^1 \mid U'_{i-}(\alpha_i) \le t \le U'_{i+}(\alpha_i) \} \tag{2.12}$$

is the interval $[U'_{i-}(\alpha_i), U'_{i+}(\alpha_i)]$.

As $f(\alpha)$ in problem (P) is a separable convex function, along any direction d the directional derivative $f'(\alpha;d)$ exists. In fact

$$f'(\alpha;d) = \max\{ <\xi, d> \mid \xi \in \partial f(\alpha) \} =$$

$$\max\{ <\xi, d> \mid U'_{i-}(\alpha_i) \le \xi_i \le U'_{i+}(\alpha_i) \},\ i = 1, \cdots, m \} \tag{2.13}$$

where ξ_i are the i-th components of vector ξ

Thus by duality theorem of linear programming, $f'(\alpha; d)$ can be expressed as a minimum value of the following dual problem

$$f'(\alpha; d) = \min \sum_{i=1}^{m} U'_{i-}(\alpha_i) v_i + \sum_{i=1}^{m} U'_{i+}(\alpha_i) w_i$$

$$\text{s.t.} \quad v_i + w_i = d_i, \; i = 1, \cdots, m$$

$$v_i \leq 0, \; w_i \geq 0, \; i = 1, \cdots, m.$$

The formula is very important in the practical calculation for the unary convex programming.

Summarizing the above results in operational terms, we have a special descent procedure that can be implemented in the unary programming.

3 Algotithm

In this section we describe a trust region method with the descent directions for the unary convex programming. Our algorithm as follows

Algorithm TR

Initialization step

Choose paramenters $0 < \eta_1 < \eta < \eta_2 < 1$, and $0 < \gamma_1 < \gamma_2 < 1 < \gamma_3$.

Select an initial trust region radius $\Delta_0 > 0$ and a maximal trust region radius $\Delta_{max} > \Delta_0$, give a starting point $\alpha_0 \in \Omega$. Set $k = 0$, go to the main step.

Main step

1. Evaluate $f_k = f(\alpha_k)$, $U'_{i-}(\alpha_i^k)$ and $U'_{i+}(\alpha_i^k)$

2. Solve the trust region subproblem as the form given by

$$(S_k) \quad \begin{aligned} \min \quad & \Phi_k(v, w) = \sum_{i=1}^{m} U'_{i-}(\alpha_i) v_i + \sum_{i=1}^{m} U'_{i+}(\alpha_i) w_i \\ \text{s.t.} \quad & v_i \leq 0, \; w_i \geq 0, \; i = 1, \cdots, m. \\ & \alpha_i^k + v + w \in \Omega \\ & \| v + w \| \leq \Delta_k. \end{aligned}$$

Obtain the solutions v^k, w^k and optimal value $\Phi_k(v^k, w^k)$.

3. If $\| \Phi_k(v^k, w^k) \| \leq \epsilon$, stop with the ϵ-approximate solution α_i^k. Otherwise go to next step.

4. Calculate the actual reduction of the objective function

$$\text{Ared}_k(d_k) = f(\alpha^k) - f(\alpha^k + d^k), \tag{3.1}$$

where

$$d^k = v^k + w^k \tag{3.2}$$

and

$$\eta_k = \frac{\text{Ared}_k(d_k)}{\Phi_k(v^k, w^k)} \tag{3.3}$$

5. If $\eta_k < \eta_1$, then $\Delta_k \leftarrow \gamma_1 \Delta_k$ and go to step 2. Otherwise, set

$$\alpha^{k+1} = \alpha^k + d^k \tag{3.4}$$

and take

$$\Delta_{k+1} = \begin{cases} \min(\gamma_3\Delta_k, \Delta_{max}) & \text{if } \eta_k \geq \eta_2 \\ \gamma_2\Delta_k & \text{if } \eta_k \leq \eta \\ \Delta_k & otherwise \end{cases} \tag{3.5}$$

let $k \leftarrow k+1$ and go to step 1.

Remark: If the l_∞ norm or the l_1 norm is used in the trust region constraint, the subproblem (S_k) simply becomes an LP problem of W decision variables v^k, w^k. After sloving it, we take $d^k = v^k + w^k$ and consider $\alpha^{k+1} = \alpha^k + d^k$ as a condidate for next iterative point, provided that $\alpha^{k+1} = \alpha^k + d^k$ can pass the test stated in step 5.

4 The Global Convergence of Algorithm TR

We first derive a global convergence of the proposed algorithm. We show that every accumulation point of $\{\alpha_k\}$ is a stationary point of f i.e.,

$$f'(\alpha; d) \geq 0 \text{ for all } d \in \Omega \tag{4.1}$$

Lemma 4.1 We have that

$$|A\,red_k(d_k) - \Phi_k(v^k, w^k)| = o(\Delta_k). \tag{4.2}$$

Proof By the definitions of $A\,red_k(d_k)$ and $\Phi_k(v^k, w^k)$, we have that

$$\begin{aligned} &|A\,red_k(d_k) - \Phi_k(v^k, w^k)| = \\ &|f(\alpha^k + d^k) - f(\alpha^k) + \Phi_k(v^k, w^k)| = \\ &|f(\alpha^k) - f(\alpha^k + d^k) + f'(\alpha^k; d^k)| = \\ &\| d^k \| = o(\Delta_k). \end{aligned} \tag{4.3}$$

□

For any α and $\Delta > 0$, let $\bar{w}$ and $\bar{v}$ be the solution of the subproblem (S_k) at point $\alpha_k = \alpha$ and $\Psi(\alpha, \Delta)$ be the minimum value, i.e.,

$$\Psi(\alpha, \Delta) = \sum_{i=1}^m U'_{i-}(\alpha_i)\bar{v}_i + \sum_{i=1}^m U'_{i+}(\alpha_i)\bar{w}_i \tag{4.4}$$

Lemma 4.2 We have that

$$\Psi(\alpha^k, \Delta_k) \leq \min\{1, \Delta_k\}\Psi(\alpha^k, 1). \tag{4.5}$$

Proof If $\Delta_k \geq 1$, it is clear to see that

$$\Psi(\alpha^k, \Delta_k) \leq \min\{1, \Delta_k\}\Psi(\alpha^k, 1). \tag{4.6}$$

On the other hand if $\Delta_k < 1$, then by the convex property of objective function $\Phi_k(v, w)$,

$$\Phi_k(\Delta_k v^k, \Delta_k w^k) \leq (1 - \Delta_k)\Phi_k(0, 0) + \Delta_k\Phi_k(v^k, w^k),$$

where v^k, w^k is the solution of the subproblem (S_k) when $\Delta_k = 1$ at α^k. Since $\Delta_k \| \overline{v^k} + \overline{w^k} \| \leq \Delta_k$, we have that $\Delta_k\overline{v^k}$ and $\Delta_k\overline{w^k}$ is a feasible solution of the subproblem (S_k) with Δ_k. Hence

$$\Psi(\alpha, \Delta) \leq \Phi(\Delta_k\overline{v^k}, \Delta_k\overline{w^k}) \leq \Delta_k\Phi_k(\overline{v^k}, \overline{w^k})) = \Delta_k\Psi(\alpha^k, 1), \tag{4.7}$$

(4.6) and (4.7) mean that (4.5) is true. □

Theorem 4.1 Let $\{\alpha^k\}$ be the sequence generated by the proposed algorithm, then we have that

$$\lim_{k\to\infty}\sup\{\Psi(\alpha^k,\Delta^k)\} = 0 \tag{4.8}$$

that is, at least one accumalation point α^* of $\{\alpha^k\}$ is a stationary point of the problem (P).

Proof Let $d^k = v^k + w^k$ be the solution of the subproblem (S_k). If $\| d^k \| \le 1$, it is clear to see that

$$\Psi(\alpha^k,\Delta^k) = \Psi(\alpha^k, \| d^k \|) \ge \Psi(\alpha_k, 1).$$

If $\| d^k \| \ge 1$, then by the convexity property of $\Phi_k(v,w)$,

$$\Phi_k\left(\frac{v^k}{\| d^k \|}, \frac{w^k}{\| d^k \|}\right) \le \left(1 - \frac{1}{\| d^k \|}\right)\Phi_k(0,0) + \frac{1}{\| d^k \|}\Phi_k(v^k,w^k).$$

From above, we have proved

$$\Psi(\alpha^k,1) \le \min\{1, \frac{1}{\| d^k \|}\}\Psi(\alpha^k,\Delta_k). \tag{4.9}$$

If the conclusion (4.8) is not true, as from $\| d^k \| \le \Delta_k \le \Delta_{max}$, we have that $\frac{1}{\| d^k \|}$ is bounded away from zero, we would have an $\epsilon > 0$ such that

$$\Psi(\alpha^k,1) \le - \epsilon, \ \forall k \tag{4.10}$$

We now prove that this is impossible. By the definition of $\Psi(\alpha^k, \Delta_k)$ and (4.10), we have

$$- \Phi_k(v^k,w^k) = - \Psi(\alpha^k,\Delta_k) \ge - \min\{1,\Delta_k\}\Psi(\alpha^k,1) \ge \min\{1,\Delta_k\}\epsilon, \tag{4.11}$$

According to (4.11), we have that

$$f(\alpha^k) - f(\alpha^{k+1}) = Ared_k(d_k) \ge - \eta_1\Phi_k(v^k,w^k) \ge \eta_1\min\{1,\Delta_k\}\epsilon \tag{4.12}$$

As the function f is bounded, from (4.12), we know that $\sum_{k=0}^{\infty}\min\{1,\Delta_k\}$ is convergent, which implies that

$$\lim_{k\to\infty}\Delta_k = 0. \tag{4.13}$$

On the other hand, from (4.2) and (4.11), we have

$$\left|\frac{Ared_k(d_k)}{\Phi_k(v^k,w^k)} - 1\right| \le \frac{o(\Delta_k)}{\min\{1,\Delta_k\}\epsilon} \quad \text{as } k\to\infty.$$

This implies $\eta_k \to 1$ as $k\to\infty$, i.e., for large k, $\eta_k \ge \eta$ and hence according to the updating rule for Δ_k, $\Delta_{k+1} \ge \Delta_k$. That is, the trust region radius will be bounded away from zero, which contradicts (4.13).

So, at least one accumulation point of α^* of $\{\alpha^k\}$ is a stationary point of the problem (P).

□

From above, we have the global convergence of the proposed algorithm.

5 Local Convergence to a Strongly Unique Solution

Let $\alpha^* \in R^m$ be a srongle unique solution of the problem (P), that is, there exist constants

$\beta > 0$, $\epsilon > 0$, such that

$$f(\alpha) - f(\alpha^*) \geq \beta \|\alpha - \alpha^*\|, \ \forall \alpha \in \Omega, \ \|\alpha - \alpha^*\| \leq \epsilon \tag{5.1}$$

We further assume that

$$\alpha \to \alpha^* \tag{5.2}$$

The condition has been introduced by Zhang in [5]. Moreover, Zhang gave the following theorem.

Theorem 5.1 If the directional derivative

$$f'(\alpha^*; d) > 0, \ \forall d \in \{d \mid \|d\| = 1\}, \ \alpha^* + d \in \Omega, \tag{5.3}$$

then the condition (5.1) must hold at α^*.

Theorem 5.2 If the sequence $\{\alpha_k\}$ generated by the proposed algorithm converges to a stationary point, then $d_k \to 0$.

Proof Let $\Omega = \{k \in N \mid \eta_k \geq \eta_1\}$ denote the set of all iteration indices belonging to accepted steps. The statement is trivial for the subsequence $\{d_k\}$ $k \in \Omega$. If, however, $N \setminus \Omega$ is infinite, even $\Delta_k \to 0$ holds. For every $k \in N \setminus \Omega$, Δ_k is reduced by the constant factor $\gamma_1 < 1$, while for $k \in \Omega$, $\Delta_{k+1} \to \Delta_k$ holds. $\Delta_k \to 0$ implies $d_k \to 0$.

From above, the conclusion of theorem holds. □

Theorem 5.3 Let the sequence $\{\alpha_k\}$ generated by the proposed algorithm convenge to a strongly unique minimizer α^* of f on Ω. Then $\{\alpha_k\}$ converges α^* superlinearly. That is,

$$\lim_{k \to \infty} \frac{\|\alpha_{k+1} - \alpha^*\|}{\|\alpha_k - \alpha^*\|} = 0. \tag{5.4}$$

Proof The above made assumptions imply that

$$f(\alpha_k + d) = f(\alpha_k) + f'(\alpha_k; d) + o(\|d\|) \tag{5.5}$$

uniformly for all k. Since $d_k \to 0$, we have that

$$\begin{aligned} \|\alpha_k + d_k - \alpha^*\| \leq \frac{1}{\beta}(f(\alpha_k + d_k) - f(\alpha^*)) = \\ \frac{1}{\beta}[f(\alpha_k) + f'(\alpha_k; \alpha^* - \alpha) - f(\alpha^*)] + o(\|d_k\|) = \\ o(\|\alpha^* - \alpha_k\|) + o(\|d_k\|) = \\ o(\|\alpha^* - \alpha_k\|). \end{aligned}$$

Therefore,

$$\|\alpha_k + d_k - \alpha^*\| = o(\|\alpha^* - \alpha_k\|) \tag{5.6}$$

Finally, we have to show that dk is accepted, i.e., $\eta_k \geq \eta_1$ for large k. From the accepting step value in the algorithm, we have that, from (4.5),

$$|\eta_k - 1| = \left| \frac{f(\alpha_k + d_k) - f(\alpha_k) - f'(\alpha_k; d_k)}{f'(\alpha_k; d_k)} \right| = \frac{|o(\|d_k\|)|}{\min\{1, \Delta_k\} |\Psi(\alpha^k; 1)|} \to 0, \quad \text{as } k \to \infty.$$

This implies that $\eta_k \geq \eta_1$ for large k which ensures that d_k is accepted for sufficiently large k. Thus the proof is complete. □

We have studied the convergence properties of the trust region method for the unary convex programming problems. We feel that the numercal test will be implemeted in practice further.

References:

[1] Clarke F H. Optimization and Nonsmooth Analysis[J]. Canadian Math Soc Series of Monographs and Advanced Texts. John Wiley, 1983.

[2] Goldfarb D, Wang S. Partial-update Newton Methods for Unary, Factorable, and Partally Separable Optimization[J]. SIAM J Optimization, 1993, 3: 382- 397

[3] MoCormick G P, Sofer A. Optimization with Unary Function[J]. Mathematical Programing, 1991, 52: 167- 178.

[4] Rockafellar R T. Monotropic Programming, Descent Algorithms and Duality [J]. Nonlinear Programming, 1972: 327- 366.

[5] Rockafellar R T. Convex Analysis [J]. Princeton Mathematics Services, Princeton University, Princeton, NJ, 1970.

[6] Zhang J. Superlinear Convergence of a Trust Region-Type Successive Linear Programming Method[J]. J Optimization Theory and Application, 1989, 61: 295- 310.

[7] Zhang J, Zhu D. A Bilevel Programming Method for Pipe Network Optimization [J]. SIAM J Optimization, 1996, 3: 838- 857.

非光滑单值优化的信赖域算法

朱德通

(上海师范大学 数学科学学院, 上海 200234)

摘要: 提供了求解非光滑单值优化问题的信赖域算法. 基于线性规划的对偶理论, 将目标函数的方向导数转化成线性规划, 从而使信赖域子问题容易数值求解. 在合理的条件下, 证明了算法的整体收敛性和局部超线性收敛速率.

关键词: 信赖域; 单值优化; 收敛性