维普资讯 http://www.cqvip.com

第30卷第2期 上海师范大学学报(自然科学版) Vol. 30, No. 2
2001年6月 J. of Shanghai Teachers Univ. (Natural Sciences) Jun. 2001

A Nonmonotone Line Search Technique for Nonsmooth Unary Optimization

FU Jun

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

Abstract: We present a nonmonotone line search 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 calculation for nonmonotone line search subproblems. A theoretical analysis proves that the proposed algorithm is globally convergent and has a local superlinear rate under some mild conditions.

Key words: nonmonotone line search technique; unary optimization; nonlinear unconstrained minimization; convergence

CLC number: O241.7 **Document code**: A **Article ID**: 1000-5137(2001)02-0034-07

1 Introduction

We consider the nonlinear unconstrained minimization problem

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

where $f(x): \mathscr{R}^n \longrightarrow \mathscr{R}^1$ is a convex function. Following GOLDFARB and WANG in [1] and MCCORMICK and SOFER in [2] 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 $i = 1, \cdots, m$, $m \geq n$, $\alpha_i(x) = a_i^T x$. a_i is a constant vector of size $n \times 1$ and $U_i(\cdot)$ is a unary convex function, i.e., $U_i(\alpha_i): \mathscr{R}^1 \rightarrow \mathscr{R}^1$, not necessarily differentiable (e.g., piecewise linear or quadratic). Many proposed algorithms for the unary optimization utilized the derivatives of unary functions involving twice continuously differentiable derivatives. However it is not always supposed to be, or even given as a finite function on all of $\mathscr{R}^1$. In such convex unary functions, there are many special ways of generalized directions of descent, and the duality of the linear programming can play a very strong role. Just by the duality idea, Zhu gave a special direction of descent for unary convex functions using the trust region strategy in [7].

Received date: 2000-11-13

Foundation item: Supported by Postgraduate Foundation of Shanghai Teachers University.

Biography: FU Jun (1977-), male, Master, College of Mathematical Sciences, Shanghai Teachers University.

维普资讯 http://www.cqvip.com

Line search technique is an important and simple class of iterative methods for solving nonlinear optimization problems to assure global convergence. In 1986, L. Grippo, etc. gave a nonmonotone line search technique in [5]. This strategy will also force global convergence of the iterates from an arbitrary starting point to a stationary point which satisfies the first-order necessary conditions.

The main purpose of this paper is to propose a nonmonotone line search method with the generalized directions of the descent by adopting the duality idea.

2 Optimality and Elementary Directions of Descent

Every unary convex programming problem can be described as

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

where Ω is a convex polyhedron in $\mathscr{R}^m$, and for $i = 1, 2, \cdots, m$, $m \geqslant n$, $a_i(x) = a_i^T x$, a_i is a constant vector of size $n \times 1$ and $U_i(\cdot)$ is a unary convex function. Each U_i has a right derivative $U'_{i-}(\alpha_i)$ and a left derivative $U'_{i-}(\alpha_i)$ at every $\alpha_i \in \mathscr{R}^1$. These are nondecreasing functions of α_i that satisfy $-\infty < U'_{i-}(\alpha_i) \leqslant U'_{i+}(\alpha_i) < +\infty$.

In problem (P), the minimal function $f(x)$ is a convex function defined on Ω. Thus x is a feasible solution 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,

$$f^0(x;d) \geqslant 0 \qquad \text{for all } d \in \Omega, \tag{2.1}$$

where $f^0(x;d)$ is the generalized directional derivative of f at x in the direction $d \in \mathscr{R}^n$. A function f is said to be regular at x if the one-sided directional derivative $f'(x;d)$ exists for all directions $d \in \mathscr{R}^n$ and

$$f'(x;d) = f^0(x;d). \tag{2.2}$$

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 of ROCKAFELLAR [3]) and a convex function always has a one-sided directional derivative. Convex analysis says that $f^0(x;d)$ and $f'(x;d)$ coincide for a convex function. 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 $f(x)$

$$f'(x;d) = \sum_{i,\, d_i > 0} U'_{i+}(a_i) d_i + \sum_{i,\, d_i < 0} U'_{i-}(a_i) d_i. \tag{2.3}$$

Because the value of $f'(x;d)$ depends on the sign of d_i, its calculation is inconvenient.

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

$$\partial f(\alpha) = \{g \in \mathscr{R}^m \mid g^T d \leqslant f^0(\alpha;d),\ \forall\, d \in \mathscr{R}^m\}, \tag{2.4}$$

and because of separability, this reduces to

$$\partial f(\alpha) = \partial U_1(\alpha_1) \times \partial U_2(\alpha_2) \times \cdots \times \partial U_m(\alpha_m). \tag{2.5}$$

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

$$\partial U_i(\alpha_i) = \{t_i \in \mathscr{R}^1 \mid U'_{i-}(\alpha_i) \leqslant t_i \leqslant U'_{i+}(\alpha_i)\} \tag{2.6}$$

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

维普资讯 http://www.cqvip.com

derivative $f'(\alpha;d)$ exists. In fact,

$$f'(\alpha;d)=\max\{<\xi,d>|\xi\in\partial f(\alpha)\}=$$
$$\max\{<\xi,d>|U'_{i-}(\alpha_i)\leqslant\xi_i\leqslant U'_{i+}(\alpha_i),\ i=1,2,\cdots,m\},\quad(2.7)$$

where ξ_i are the 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 (see [7])

$$f'(\alpha;d)=\min\ \sum_{i=1}^{m}U'_{i-}(\alpha_i)\nu_i+\sum_{i=1}^{m}U'_{i+}(\alpha_i)\omega_i$$
$$s.t.\ \nu_i+\omega_i=d_i,\quad i=1,2,\cdots,m.$$
$$\nu_i\leqslant 0,\ \omega_i\geqslant 0,\ i=1,2,\cdots,m.$$

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

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

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

3 Nonmonotone Line Search Algorithm

We describe a nonmonotone line search method with the descent directions for the unary convex programming. Our algorithm is as follows:

Algorithm NLS

Initialization Step

Date: $\alpha_0\in\Omega$, integer $M\geqslant 0$, $\gamma\in(0,1)$, $\sigma\in(0,1)$, $\varepsilon>0$.

Main Step:

Step 1: Set $k=0$, $m(0)=0$ and compute $f_0=f(x_0)$, $U'_{i-}(\alpha_i^0)$ and $U'_{i+}(\alpha_i^0)$.

Step 2: Solve the NLS subproblem as the form given by

$$(S_k)\quad \min\ \Phi_k(\nu,\omega)=\sum_{i=1}^{m}U'_{i-}(\alpha_i)\nu_i+\sum_{i=1}^{m}U'_{i+}(\alpha_i)\omega_i$$
$$\text{s.t. } \nu_i\leqslant 0,\ \omega_i\geqslant 0,\ i=1,2,\cdots,m.\quad(3.1)$$
$$\alpha^k+\nu+\omega\in\Omega.$$
$$\|\nu+\omega\|_\infty\leqslant 1.$$

Obtain the solution ν^k,ω^k and the optimal value $\Phi_k(\nu^k,\omega^k)$.

Step 3: If $|\Phi_k(\nu^k,\omega^k)|\leqslant\varepsilon$, stop with the approximate solution α^* , otherwise go to the next step.

Step 4: Set $\lambda=1$, compute $d_k=\nu^k+\omega^k$.

Step 5: Compute $f(\alpha_k+\lambda d_k)$. If

$$f(\alpha_k+\lambda d_k)\leqslant\max_{0\leqslant j\leqslant m(k)}[f(\alpha_{k-j})]+\gamma\lambda f'(\alpha_k;d_k)\quad(3.2)$$

holds, set $\alpha_{k+1}=\alpha_k+\lambda d_k$, $k=k+1$, $f(\alpha_{k+1})=f(\alpha_k+\lambda d_k)$, $m(k)\leqslant\min[m(k-1)+1,M]$ and go to step 2; otherwise go to the next step.

Step 6: Set $\lambda=\sigma\lambda$ and go to Step 5.

Remark: If the l_∞ norm or the l_1 norm is used in the constraint, the subproblem (S_k) simply becomes an LP problem of ν^k,ω^k. After sloving it, we take $d_k=\nu^k+\omega^k$ and consider $\alpha_{k+1}=\alpha_k+d_k$ as a candidate for the next iterative point, provided that $\alpha_{k+1}=\alpha_k+d_k$ can pass the test stated in Step 5.

维普资讯 http://www.cqvip.com

4 The Global Convergence of Algorithm NLS

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

$$\Psi(\alpha,\Delta) = \sum_{i=1}^{m} U'_{i-}(\alpha_i)\bar{\nu}_i + \sum_{i=1}^{m} U'_{i+}(\alpha_i)\bar{\omega}_i. \tag{4.1}$$

Lemma 4.1 $\Psi(\alpha,\Delta)$ is a monotonically decreasing function of Δ, that is, if $0 \leqslant \Delta_1 \leqslant \Delta_2$, then $\Psi(\alpha,\Delta_1) \geqslant \Psi(\alpha,\Delta_2)$.

Lemma 4.2[7] $\Psi(\alpha^k,\Delta) \leqslant \min\{1,\Delta\}\Psi(\alpha^k,1)$.

Lemma 4.3 $\limsup_{k\to\infty}\{\Psi(\alpha^k,1)\} = 0$ if and only if $\limsup_{k\to\infty}\{\Psi(\alpha^k,\Delta)\} = 0$, $\forall\ \Delta \geqslant 0$.

Proof If $\limsup_{k\to\infty}\{\Psi(\alpha^k,\Delta)\} = 0$, $\forall\ \Delta \geqslant 0$ holds, it is obvious that $\limsup_{k\to\infty}\{\Psi(\alpha^k,1)\} = 0$.

Assume that $\limsup_{k\to\infty}\{\Psi(\alpha^k,1)\} = 0$ holds. Obviously, the Lemma is true for $\Delta = 0$.

If $0 < \Delta \leqslant 1$, then by lemma 4.2, we have $\Psi(\alpha^k,\Delta) \leqslant \Delta\Psi(\alpha^k,1)$. At the same time, by lemma 4.1, $\Psi(\alpha^k,\Delta) \geqslant \Psi(\alpha^k,1)$ always holds for $0 < \Delta \leqslant 1$. So we have $\Psi(\alpha^k,1) \leqslant \Psi(\alpha^k,\Delta) \leqslant \Delta\Psi(\alpha^k,1)$.

If $\Delta > 1$, then $\Psi(\alpha^k,\Delta) \leqslant \Psi(\alpha^k,1)$ holds. On the other hand, by the convex property of objective function $\Phi_k(\nu,\omega)$, $\Phi_k(\frac{\nu^k}{\Delta},\frac{\omega^k}{\Delta}) \leqslant (1-\frac{1}{\Delta})\Phi_k(0,0) + \frac{1}{\Delta}\Phi_k(\nu^k,\omega^k)$ $(0 < \frac{1}{\Delta} \leqslant 1)$, where ν^k, ω^k is the optimal solution of the subproblem (S_k) with Δ at α^k and $(0,0)$ is always the feasible solution of (S_k). Since $\frac{\| \nu^k + \omega^k \|}{\Delta} \leqslant 1$, we have that $\frac{\nu^k}{\Delta}$ and $\frac{\omega^k}{\Delta}$ is the feasible solution of the subproblem (S_k) with $\Delta = 1$. Hence $\Delta\Psi(\alpha^k,1) \leqslant \Phi_k(\frac{\nu^k}{\Delta},\frac{\omega^k}{\Delta}) \leqslant \Psi(\alpha^k,\Delta) = \Psi(\alpha^k,\Delta)$, that is $\Delta\Psi(\alpha^k,1) \leqslant \Psi(\alpha^k,\Delta) \leqslant \Psi(\alpha^k,1)$. The proof is completed. □

If $\limsup_{k\to\infty}\{\Psi(\alpha^k,1)\} \neq 0$, there exists an $\varepsilon > 0$, such that $\Psi(\alpha^k,1) < -\varepsilon$, since $\Psi(\alpha^k,1)$ is always less than zero. By Lemma 4.2, we have $\Psi(\alpha^k,\Delta) \leqslant \min\{1,\Delta\}\Psi(\alpha^k,1) < -\varepsilon\Delta$, for $0 < \Delta \leqslant 1$. Because of $f'(\alpha_k;d_k) \triangleq \Psi(\alpha^k, \| d_k \|)$ for any $\lambda > 0$, we have $\lambda f'(\alpha_k;d_k) \leqslant -\varepsilon\lambda \| d_k \|$, for $\| d_k \| \leqslant 1$. By Lemma 4.3, we only need to solve (S_k) with $\Delta = 1$. Now we give the global convergence theorem of the algorithm NLS.

Theorem Let $\{\alpha_k\}$ be a sequence defined by $\alpha_{k+1} = \alpha_k + \lambda_k d_k$, $d_k \neq 0$. Let $\sigma \in (0,1)$, $\gamma \in (0,1)$ and let M be a nonnegative integer. Assume that:

(i) the level set $\Omega_0 \triangleq \{\alpha : f(\alpha) \leqslant f(\alpha_0)\}$ is compact, $f(\alpha)$ is uniformly continuous on Ω_0;

(ii) there exist positive numbers $\lambda > 0$ and $\varepsilon > 0$, such that

$$\lambda f'(\alpha_k;d_k) \leqslant -\varepsilon\lambda \| d_k \|. \tag{4.2}$$

(iii) $\lambda = \sigma^{h_k}$, where h_k is the first nonnegative integer h for which

$$f(\alpha_k + \sigma^h d_k) \leqslant \max_{0\leqslant j\leqslant m(k)} [f(\alpha_{k-j})] + \gamma\sigma^h f'(\alpha_k;d_k), \tag{4.3}$$

where $m(0) = 0$ and $0 \leqslant m(k) \leqslant \min[m(k-1)+1, M]$, $k \geqslant 1$. Then the sequence $\{\alpha_k\}$ remains in Ω_0 and every limit point $\bar{\alpha}$ satisfies $f'(\bar{\alpha};d) = 0$.

Proof Let $l(k)$ be an integer such that $k - m(k) \leqslant l(k) \leqslant k$, $f(\alpha_{l(k)}) = \max_{0\leqslant j\leqslant m(k)} [f(\alpha_{k-j})]$. By (iii) we get that the sequence $\{f(\alpha_{l(k)})\}$ is nonincreasing. Now, since $f(\alpha_k) \leqslant f(\alpha_0)$ for all k, $\{\alpha_k\} \subset \Omega_0$, so $f(\alpha_{l(k)})$ admits a limit for $k \to \infty$. Moreover we obtain from (4.3) that for $k > M$:

维普资讯 http://www.cqvip.com

$$f(\alpha_{l(k)}) \leqslant f(\alpha_{l(l(k)-1)}) + \gamma\lambda_{l(k)-1} f'(\alpha_{l(k)-1}; d_{l(k)-1}). \tag{4.4}$$

Also since $\lambda_{l(k)-1} > 0$ and $f'(\alpha_{l(k)-1}; d_{l(k)-1}) \leqslant 0$, it follows from (4.4) that

$$\lim_{k\to\infty} \lambda_{l(k)-1} f'(\alpha_{l(k)-1}; d_{l(k)-1}) = 0. \tag{4.5}$$

By (ii) we have $\lambda f'(\alpha_k; d_k) \leqslant -\varepsilon\lambda \| d_k \|$, for all k , and thus, we have

$$0 = \lim_{k\to\infty} \lambda_{l(k)-1} f'(\alpha_{l(k)-1}; d_{l(k)-1}) \leqslant \lim_{k\to\infty} (-\varepsilon\lambda_{l(k)-1}) \| d_{l(k)-1} \| \leqslant 0.$$

That is to say (4.5) implies

$$\lim_{k\to\infty} \lambda_{l(k)-1} \| d_{l(k)-1} \| = 0. \tag{4.6}$$

We prove now that $\lim_{k\to\infty} \lambda_k \| d_k \| = 0$. Let $\hat{l}(k) \triangleq l(k+M+2)$. First we show, by induction, that for any given $j \geqslant 1$,

$$\lim_{k\to\infty} \lambda_{\hat{l}(k)-j} \| d_{\hat{l}(k)-j} \| = 0. \tag{4.7}$$

and

$$\lim_{k\to\infty} f(\alpha_{\hat{l}(k)-j}) = \lim_{k\to\infty} f(\alpha_{l(k)}). \tag{4.8}$$

(Here and in the sequel we assume, without loss of generality, that the iteration index k is large enough to avoid the occurrence of negative subscripts, that is $k \geqslant j-1$). For $j=1$, since $\{\hat{l}(k)\} \subset \{l(k)\}$ (4.7) follows from (4.6) , and this in turn implies $\| \alpha_{\hat{l}(k)} - \alpha_{\hat{l}(k)-1} \| \to 0$, so (4.8) holds for $j=1$, since $f(\alpha)$ is uniformly continuous on Ω_0 . Assume now that (4.7) and (4.8) hold for a given j . Then by (4.3) one can write: $f(\alpha_{\hat{l}(k)-j}) \leqslant f(\alpha_{l(\hat{l}(k)-j-1)}) + \gamma\lambda_{\hat{l}(k)-j-1} f'(\alpha_{\hat{l}(k)-j-1}; d_{\hat{l}(k)-j-1})$. Taking limits as $k \to \infty$, we have, by (4.8):

$$\lim_{k\to\infty} \lambda_{\hat{l}(k)-(j+1)} f'(\alpha_{\hat{l}(k)-(j+1)}; d_{\hat{l}(k)-(j+1)}) = 0. \tag{4.9}$$

Using the same arguments employed for deriving (4.6) from (4.5), we obtain

$$\lim_{k\to\infty} \lambda_{\hat{l}(k)-(j+1)} \| d_{\hat{l}(k)-(j+1)} \| = 0.$$

Moreover this implies $\| \alpha_{\hat{l}(k)-j} - \alpha_{\hat{l}(k)-(j+1)} \| \to 0$, by (4.8) and the uniform continuity of f on Ω_0 :

$$\lim_{k\to\infty} f(\alpha_{\hat{l}(k)-(j+1)}) = \lim_{k\to\infty} f(\alpha_{\hat{l}(k)-j}) = \lim_{k\to\infty} f(\alpha_{l(k)}).$$

Thus we conclude that (4.7) and (4.8) hold for any given $j \geqslant 1$. Now for any k ,

$$\alpha_{k+1} = \alpha_{\hat{l}(k)} - \sum_{j=1}^{\hat{l}(k)-k-1} \lambda_{\hat{l}(k)-j} d_{\hat{l}(k)-j}, \tag{4.10}$$

by the definition of $\hat{l}(k)$, we have $\hat{l}(k) - l - 1 = l(k+M+2) - k - 1 \leqslant M+1$. So (4.10) implies

$$\lim_{k\to\infty} \| \alpha_{k+1} - \alpha_{\hat{l}(k)} \| = \lim_{k\to\infty} \| \sum_{j=1}^{\hat{l}(k)-k-1} \lambda_{\hat{l}(k)-j} d_{\hat{l}(k)-j} \| = 0. \tag{4.11}$$

Since $\{f(\alpha_{l(k)})\}$ admits a limit, it follows from the uniform continuity of f on Ω_0 that

$$\lim_{k\to\infty} f(\alpha_{k+1}) = \lim_{k\to\infty} f(\alpha_{l(k)}). \tag{4.12}$$

By (4.3) we have

$$f(\alpha_{k+1}) \leqslant f(\alpha_{l(k)}) + \gamma\lambda_k f'(\alpha_k; d_k).$$

Taking limits as $k \to \infty$, by (4.12) we obtain:

$$\lim_{k\to\infty} \lambda_k f'(\alpha_k; d_k) = 0, \tag{4.13}$$

which implies, as noted before:

$$\lim_{k\to\infty} \lambda_k \| d_k \| = 0. \tag{4.14}$$

Now let $\bar{\alpha}$ be any limit point of $\{\alpha_k\}$ and relabel $\{\alpha_k\}$ a subsequence converging to $\bar{\alpha}$. Then by (4.13) either $\lim_{k\to\infty} f'(\alpha_k; d_k) = 0$, or there exists a subsequence $\{\alpha_k\}_K \subset \{\alpha_k\}$ such that: $\lim_{k\to\infty, k\in K} \lambda_k = 0$.

Because we assume that $\lim_{k\to\infty} \sup\{\Psi(\alpha^k, 1)\} \neq 0$, the first case doesn't occur. In the next case, by

维普资讯 http://www.cqvip.com

assumption (iii) there exists an index $\bar{k}$ such that for all $k \geqslant \bar{k}$, $k \in K$,

$$f(\alpha_k + \frac{\lambda_k}{\sigma} d_k) > \max_{0 \leqslant j \leqslant m(k)} [f(\alpha_{k-j})] + \gamma \frac{\lambda_k}{\sigma} f'(\alpha_k; d_k) \geqslant f(\alpha_k) + \gamma \frac{\lambda_k}{\sigma} f'(\alpha_k; d_k),$$

that is

$$\frac{f(\alpha_k + \frac{\lambda_k}{\sigma} d_k) - f(\alpha_k)}{\frac{\lambda_k}{\sigma}} > \gamma f'(\alpha_k; d_k). \tag{4.15}$$

Let now $\{\alpha_k\}_{K_1} \subset \{\alpha_k\}_K$ be a subsequence such that

$$\lim_{k \to \infty, k \in K_1} \alpha_k = \bar{\alpha}, \qquad \lim_{k \to \infty, k \in K_1} d_k = \bar{d}.$$

By (4.15) and the definition of the directional derivative of f, taking limits for $k \to \infty$, $k \in K_1$, we obtain: $f'(\bar{\alpha}; \bar{d}) \geqslant \gamma f'(\bar{\alpha}; \bar{d})$, that is $(1-\gamma) f'(\bar{\alpha}; \bar{d}) \geqslant 0$. Since $1 - \gamma > 0$ and $f'(\bar{\alpha}; \bar{d}) \leqslant 0$ for all k, we have $f'(\bar{\alpha}; \bar{d}) = 0$, and this completes the proof of the theorem. □

5 Local Convergence to a Strongly Unique Solution

In this section, the following hypotheses are required.

Hypothesis H1 $\alpha_k \to \alpha^*$.

Hypothesis H2 $\|d_k\| \to 0$.

Hypothesis H3 There exist $\theta > 0$ and $\delta > 0$ such that for any α in the ball $N(\alpha^*; \delta) = \{\alpha \mid \|\alpha - \alpha^*\| \leqslant \delta\}$, we have $f(\alpha) - f(\alpha^*) \leqslant \theta \|\alpha - \alpha^*\|$. f is said to satisfy the Lipschitz condition at α^*.

Hypothesis H4 There exist $\beta > 0$ and $\delta > 0$ such that for any α in the ball $N(\alpha^*; \delta) = \{\alpha \mid \|\alpha - \alpha^*\| \leqslant \delta\}$, we have $f(\alpha) - f(\alpha^*) \geqslant \delta \|\alpha - \alpha^*\|$.

Hypothesis H4 is called a growth condition by SACHS[6]. The condition also has been introduced by ZHANG[4]. Moreover, Zhang gave the following theorem:

Theorem 5.1 If the directional derivative $f(\alpha^*; d) > 0$, $\forall\ d \in \{d \mid \|d\| = 1\}$, $\alpha^* + d \in \Omega$, then the growth condition must hold at α^*.

Theorem 5.2 Let the sequence $\{\alpha_k\}$ generated by the proposed algorithm satisfy H1, H2, H3, and H4. Then $\{\alpha_k\}$ converges to α^* superlinearly. That is $\lim_{k \to \infty} \frac{\|\alpha_{k+1} - \alpha^*\|}{\|\alpha_k - \alpha^*\|} = 0$.

Proof The above-made assumptions imply that $f(\alpha_k + d) = f(\alpha_k) + f'(\alpha_k; d) + o(\|d\|)$ uniformly for all k. Since $\|d_k\| \to 0$, we have that

$$\|\alpha_k + d_k - \alpha^*\| \leqslant \frac{1}{\beta}(f(\alpha_k + d_k) - f(\alpha^*)) =$$

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

$$o(\|\alpha^* - \alpha_k\|) + o(\|d_k\|) = o(\|\alpha^* - \alpha_k\|).$$

Therefore $\|\alpha_k + d_k - \alpha^*\| = o(\|\alpha^* - \alpha_k\|)$.

Finally, we show that d_k is accepted and λ_i equals to 1 for sufficiently large k.

By H3, there exist $\theta > 0$ and $\delta_1 > 0$ such that for any α in the ball $N(\alpha^*; \delta_1) = \{\alpha \mid \|\alpha - \alpha^*\| \leqslant \|\delta_1\|$ we have $f(\alpha) - f(\alpha^*) \leqslant \theta \|\alpha - \alpha^*\|$.

By H4, there exist $\beta > 0$ and $\delta_2 > 0$ such that for any α in the ball $N(\alpha^*; \delta_2) = \{\alpha \mid \|\alpha - \alpha^*\| \leqslant \delta_2\}$ we have $f(\alpha) - f(\alpha^*) \leqslant \beta \|\alpha - \alpha^*\|$.

维普资讯 http://www.cqvip.com

Choosing $\hat{\delta} = \min\{\hat{\delta}_1, \hat{\delta}_2\}$, we have $\beta\|\alpha - \alpha^*\| \leqslant \|f(\alpha) - f(\alpha^*)\| \leqslant \theta\|\alpha - \alpha^*\|$. That is to say $\|f(\alpha) - f(\alpha^*)\| = O(\|\alpha - \alpha^*\|)$. For sufficiently large k_1, if $k > k_1$, $\alpha_k \in N(\alpha^*; \hat{\delta}_1) \cap N(\alpha^*; \hat{\delta}_2)$ holds $\frac{\max_{0\leqslant j\leqslant m(k)}[f_{k-j}] - f(\alpha^*)}{-f'(\alpha_k; \alpha^* - \alpha_k)} \geqslant \frac{f(\alpha_k) - f(\alpha^*)}{f(\alpha_k) - f(\alpha^*) + o(\|\alpha^* - \alpha_k\|)} = 1 - \frac{o(\|\alpha^* - \alpha_k\|)}{O(\|\alpha^* - \alpha_k\|)}$. For $\frac{o(\|\alpha^* - \alpha_k\|)}{O(\|\alpha^* - \alpha_k\|)}$, there exist $k_2 > 0$ such that $k > k_2$, $\frac{o(\|\alpha^* - \alpha_k\|)}{O(\|\alpha^* - \alpha_k\|)} \leqslant 1 - \gamma$ holds. So

$$\frac{\max_{0\leqslant j\leqslant m(k)}[f_{k-j}] - f(\alpha^*)}{-f'(\alpha_k; \alpha^* - \alpha_k)} \geqslant \gamma$$

holds for $k \geqslant \max\{k_1, k_2\}$. Hence Step 5 is satisfied and we set $\lambda_k = 1$. □

We have studied the convergence properties of the nonmonotone line search method for the unary convex programming problems and given a worked example in the appendix. We expect that the numerical test will be implemented in practice.

Acknowledgements The author would like to thank Prof. Zhu Detong for his inspiration in working work on this problem as well as many fruitful discussions with me in the early stage of this work and for his critical reading of the earlier versions of this paper.

References:

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

[2] MOCORMICK G P, SOFER A. Optimization with Unary Function[J]. Mathematical Programming, 1991, 52: 167-178.

[3] ROCKFELLAR R T. Convex Analysis, Princeton, New Jersey[M]. Princeton University Press. 1977.

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

[5] GRIPPO L, LAMPARIELLO F, LUCIDI S. A Nonmonotone Line Search Technique for Newton's Methods [J]. SIAM J. Numer Anal, 1986, 23: 707-716.

[6] SACH E. Global Convergence of Quasi-Newton-Type Algorithms For Some Nonsmooth Optimization Problems [J]. J Optimization Theory and Applications, 1983, 40: 201-219.

[7] ZHU D T. A Trust Region Algorithm for Nonsmooth Unary Optimization[J]. Journal of Shanghai Teachers University, 1999, 28: 1-9.

非光滑单值优化的非单调线搜索方法

傅 军

(上海师范大学 数学科学学院, 上海 200234)

摘要: 提供了非光滑单值优化的非单调线搜索方法. 基于线性规划的对偶定理, 目标函数的方向导数可以表示成线性规划问题, 这在实际计算非单调线搜索子问题时是非常重要的. 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: 非单调线搜索技术; 单值优化; 非线性无约束极小化; 收敛性