Improved Non-Monotone line Search Slackness technique to Find the Actual Maximum Value of any given Function

V.S.Akilandeswari

Assistant Professor. Saranathan College of Engineering,Panjappur.

e-mail:akilaharish22@gmail.com

Abstract The main aim of this paper is to introduce a new non monotone line search slackness technique to find the accurate maximum value of any function. It also deals with the convergence of the method. Then the corresponding algorithm is established along with the illustrations to prove the same.

Keywords Unconstrained optimization problem, Improved Nonmonotone line search slackness technique, global convergence, BFGS algorithm.

1.Introduction

Consider the following unconstrained optimization problem:

$$\max_{x\in\Re^n} \phi(x)$$

where $\phi(x)$ is a continuously differentiable function from R^n to R.

To find the maximum of the given function, first the function is converted to minimum function $f(x)$ by using the relation min - $\phi(x)$ =- max $\phi(x)=f(x)$

At current iteration x_k, if $g_k=\nabla f(x_k)\neq 0$, a line search method defines a search direction d_k in some way, finds a step-length by carrying some line search along d_k. Among the most popular line search rules are the Armijo rule, the Goldstein rule and the wolfe rule. The step-length α_k is found by carrying some line search along the direction d_k, and then obtain the next iteration as

$$x_{k+1} = x_k + \alpha_k d_k$$

The search direction d_k can be determined by many methods.

The traditional line searches require the function value descent monotonically at every iteration, namely:

$$f(x_{k+1}) \leq f(x_k)$$

$\alpha_k \geq 0$ is obtained by the line search.

Grippo et al introduced a highly innovative method called the nonmonotone line search technique which *is* to determine the step-length α_k such that

$$f(x_k+\alpha_k d_k) \leq \max_{0\leq j\leq m(k)}\{f(x_{k-j}) - \sigma \propto_k (-g_k^T d_k)$$

where {m(k)} is an integer sequence satisfying the following conditions:

m(0)=0, and 0$\leq m(k) \leq \min\{m(k-1)+1, M\}$

for some positive integer MϵZ_+.

Sun et al. put forward the so-called nonmonotone F-rule, which is to determine $\propto_k$ such that

$$f(x_{k+1}) = f(x_k + \alpha_k d_k) \leq \max\left\{f(x_k), \sum_{r=0}^{m(k)-1} \lambda_{kr} f(x_{k-r})\right\} - \sigma(t_k)$$

Where $\sum_{r=0}^{m(k)-1} \lambda_{kr} = 1.$

Ping Hu, Xu qing Liu introduced nonmonotone slackness rule

$$f(x_{k+1}) = f(x_k + \alpha_k d_k) \leq \mu_k f_{l(k)} - \sigma_k$$

where $\sigma_k = \sigma(t_k)$, $f_{l(k)} = \{\max_{0\leq r\leq m(k)} f(x_{k-r})\}$ and $\mu_k = \begin{cases} \lambda^{h_k} & if \;\; f_{l(k)} > 0 \\ \lambda^{-h_k} & if \;\; f_{l(k)} \leq 0 \end{cases}$

with $\lambda \geq 1, h_k \geq 1$ and $\sum_{k=0}^{\infty} h_k = \gamma$, in which γ is a finite constant, m(k)=min{k,M-1}, M≥1 is a positive integer.

In this paper, a new method is introduced to find the maximum of a given function. The paper is arranged as follows. In section 2, a new non-monotone rule is introduced followed by proving its convergence using lemmas and theorem. In section 3, The numerical tests are established to illustrate the effectiveness of the algorithm.

2. Improved Nonmonotone line search slackness technique

2.1 Proof for Global Convergence

2.1.1 Assumptions

A1: Assume that the function f: $R^n \to R$ is bounded and differentiable on the level set

£c:={ $x \in R^n : f(x) \leq c\,|\,f(x_0)\,|$ } for a given constant c ≥ 1.

A2: The search direction d_k is determined by BFGS algorithm:

$$d_k = \begin{cases} -g_k & if \;\; g_k^T d_k \geq 0 \\ -H_k g_k & otherwise \end{cases}$$

where $g_k \neq 0$, H_k is generated by the BFGS formula.

2.1.2 Definitions

D1: The function $\sigma: [0, +\infty] \to [0, +\infty]$ is a forcing function (F-function) if for any sequence

$\{t_i\} \subset [0, +\infty]$

$\lim_{i\to\infty} \sigma(t_i) = 0$ implies $\lim_{i\to\infty} t_i = 0$.

D2: Let $\eta = \sup\{||g(x) - g(y)||: x, y \in £\} > 0$. Then the mapping

$\delta:[0,+\infty]$ $[0,+\infty]$ defined by

$$\delta(t)=\begin{cases}\inf\left\{\|x-y\|\,\middle|\,\|g(x)-g(y)\|\geq t\right. & t\in[0,\eta)\\ \lim\limits_{\delta\to\eta^-}\delta(s) & t\in[\eta,+\infty]\end{cases}$$

is the reverse modulus of continuity of gradient g(x).

2.1.3 *Improved Nonmonotone Line Search Slackness Technique*

Let M be a non-negative integer. For each k. m(k) satisfy

m(0)=0 and m(k)=min{k,M-1} $k\geq 1$.

The main idea of this rule is to determine the bounded step length α_k along the direction $\mathrm{d_k}$ such that

$$f_{k+1}\leq\mu_k f_{l(k)}-\sigma_k \qquad -(1)$$

where $\sigma_k=\sigma(t_k)$, is a forcing function and $t_k=-g_k^T d_k/\|d_k\|$

$$f_{l(k)}=\{\max_{0\leq r\leq m(k)} f(x_{k-r})\}\quad \mu_k=\begin{cases}e^{\lambda h_k} & if\quad f_{l(k)}>0\\ e^{-\lambda h_k} & if\quad f_{l(k)}\leq 0\end{cases}$$

with $\lambda\geq 0, h_k\geq 1$ and $\sum_{k=0}^{\infty}h_k=\gamma$ in which γ is a finite constant. Set $x_{k+1}=x_k+\alpha_k d_k$.

2.1.4 Lemmas and Theorems

L1: *The inequality*

$$f_{k+1}\leq|f_0|e^{\lambda s(k)}-\sigma_k$$

holds for k=0, 1, 2, …… if α_k *satisfies (1), where s(k):=*$\sum_{i=0}^{k}h_i$.

Proof: The principle of mathematical induction will be used to prove the conclusion.

Assume k=0

$$\begin{aligned}f_1&\leq\mu_0 f_{l(0)}-\sigma_0\\ &=\mu_0 f_0-\sigma_0 \qquad (\ f_{l(k)}=\{\max_{0\leq r\leq m(k)} f(x_{k-r})\})\\ f_1&\leq e^{\lambda h_0}f_0-\sigma_0\end{aligned}$$

Since $e^{-\lambda h_k}\leq\mu_k\leq e^{\lambda h_k}$.

$$f_1 \le e^{\lambda s(0)} | f_0 | - \sigma_0$$

Hence the lemma is true when k=0.

Assume that L1 is true for all values till k. Thus

$$f_k \le | f_0 | e^{\lambda s(k-1)} - \sigma_{k-1}.$$

Now to prove for k+1.

It is known that

$$\begin{aligned} f_{k+1} &\le \mu_k f_{l(k)} - \sigma_k \\ &\le \mu_k \{| f_0 | e^{\lambda s(l(k)-1)} - \sigma_{l(k)-1}\} - \sigma_k \\ &\le \mu_k | f_0 | e^{\lambda s(l(k)-1)} - \sigma_k \\ &\le e^{\lambda h_k} | f_0 | e^{\lambda s(l(k)-1)} - \sigma_k \\ &= | f_0 | e^{\lambda s(k)} - \sigma_k \end{aligned}$$

Hence L1 is true for k+1.

Thus $f_{k+1} \le | f_0 | e^{\lambda s(k)} - \sigma_k$ is true for all k.

From $\begin{aligned} f_{k+1} &\le \mu_k f_{l(k)} - \sigma_k \\ &\le \mu_k f_{l(k)} \end{aligned}$

with $l(k) \le k$.

Therefore $\begin{array}{lll} \mu_k \ge 1 & if & f_{l(k)} > 0 \\ 0 \le \mu_k \le 1 & if & f_{l(k)} \le 0 \end{array}$

Hence $f(x_k) \le f(x_{l(k)}) \le f(x_{k+1})$.

From

$$\begin{aligned} f_{k+1} &\le | f_0 | e^{\lambda s(k)} - \sigma_k \\ &\le e^{\lambda \gamma} | f_0 | \end{aligned}$$

$$also \quad \lambda\gamma \ge 0 \Rightarrow f_{k+1} > f_0$$

Take $h_k=(k+1)^{-p}$ with p>1.

$$\sum_{k=0}^{\infty} h_k = \sum_{k=0}^{\infty} (k+1)^{-p} = \sum_{n=0}^{\infty} \frac{1}{n^p}$$

Thus $\sum_{k=0}^{\infty} h_k$ converges to γ

Therefore $e^{\lambda \sum_{k=0}^{\infty} h_k} = e^{\lambda\gamma} = c$

L2: *Assume $\{x_k\}$ is a sequence generated according to the search direction d_k satisfying A2 and the step-length α_k determined by (1). Then $\{x_k\} \subseteq £_c$.*

Proof: By L1, we obtain

$$f_{k+1} \le |f_0| e^{\lambda s(k)} - \sigma_k \le e^{\lambda\gamma} |f_0| = c|f_0|$$

where c=$e^{\lambda\gamma} \ge 1$. Then, by the definition of $£_c$, $\{x_k\} \subseteq £_c$.

L3: $f_{jM+r} \le |f_0| e^{\lambda s(jM+r-1)} - e^{-\lambda s(jM+r-1)} \sum_{i=0}^{j} \sigma_{t(i)}$ *holds for j=0,1,2,……n*

Where $1 \le r \le M$ *and* $t(i) \in \{iM + r - 1 / 1 \le r \le M\}$

Proof: It is known that

$$f_{k+1} \le |f_0| e^{\lambda s(k)} - \sigma_k$$
$$f_r \le |f_0| e^{\lambda s(r-1)} - \sigma_{r-1}$$
$$\le |f_0| e^{\lambda s(r-1)} - e^{-\lambda s(r-1)} \sigma_{r-1}$$
$$f_r \le |f_0| e^{\lambda s(r-1)} - e^{-\lambda s(r-1)} \sigma_{t(0)}$$

where $\sigma_{t(0)} = \min_{1 \le r \le M} \sigma_{r-1}$. Hence L3 is true for j=0.

Assume L3 is true for j=k.

$$f_{kM+r} \le |f_0| e^{\lambda s(kM+r-1)} - e^{-\lambda s(kM+r-1)} \sum_{i=0}^{k} \sigma_{t(i)}$$

Now to prove for j=k+1.

To Prove: $f_{(k+1)M+r} \le |f_0| e^{\lambda s((k+1)M+r-1)} - e^{-\lambda s((k+1)M+r-1)} \sum_{i=0}^{k+1} \sigma_{t(i)}$ -(*)

While proving for j=k+1, it is also proved for all r.

Note that $(k+1)M+1 \le jM+r \le (k+2)M$
$m(jM+r) = M-1$

When r=1

$$f_{(k+1)M+1} \le \mu_{(k+1)M} f_{l((k+1)M)} - \sigma_{(k+1)M}$$
$$= \mu_{(k+1)M} \max\{ f_{(k+1)M}, f_{(k+1)M-1}, \dots\dots\dots, f_{kM+1}\} - \sigma_{(k+1)M}$$

From

$$kM+1 \le l((k+1)M) \le (k+1)M$$

$$f_{kM+r} \le | f_0 | e^{\lambda s(kM+r-1)} - e^{-\lambda s(kM+r-1)} \sum_{i=0}^{k} \sigma_{t(i)}$$

$$f_{l((k+1)M)} \le | f_0 | e^{\lambda s(l((k+1)M)-1)} - e^{-\lambda s(l(k+1)M)-1)} \sum_{i=0}^{(l(k+1)M-r)/M} \sigma_{t(i)}$$

$$f_{l((k+1)M)} \le | f_0 | e^{\lambda s(l((k+1)M)-1)} - e^{-\lambda s(l(k+1)M)-1)} \sum_{i=0}^{k} \sigma_{t(i)}$$

$$f_{(k+1)M+1} \le \mu_{(k+1)M} \{| f_0 | e^{\lambda s(l((k+1)M)-1)} - e^{-\lambda s(l(k+1)M)-1)} \sum_{i=0}^{k} \sigma_{t(i)}\} - \sigma_{(k+1)M}$$

$$\le e^{\lambda h_{(k+1)M}} | f_0 | e^{\lambda s(l((k+1)M)-1)} - e^{-\lambda h_{(k+1)M}} e^{-\lambda s(l(k+1)M)-1)} \sum_{i=0}^{k} \sigma_{t(i)} - \sigma_{t(k+1)}$$

$$\le e^{\lambda h_{(k+1)M}} | f_0 | e^{\lambda s(((k+1)M)-1)} - e^{-\lambda h_{(k+1)M}} e^{-\lambda s((k+1)M)-1)} \sum_{i=0}^{k} \sigma_{t(i)} - \sigma_{t(k+1)}$$

$$\le | f_0 | e^{\lambda s((k+1)M)} - e^{-\lambda s((k+1)M))} \sum_{i=0}^{k} \sigma_{t(i)} - e^{-\lambda s((k+1)M))} \sigma_{t(k+1)}$$

$$f_{(k+1)M+1} \le | f_0 | e^{\lambda s((k+1)M)} - e^{-\lambda s((k+1)M))} \sum_{i=0}^{k+1} \sigma_{t(i)}$$

Hence (*) is true for r=1.

Assume (*) is true for r=r_0.

$$f_{(k+1)M+r_0} \le | f_0 | e^{\lambda s((k+1)M+r_0-1)} - e^{-\lambda s((k+1)M+r_0-1)} \sum_{i=0}^{k+1} \sigma_{t(i)}$$

When r=r_0+1

$$f_{(k+1)M+r_0+1} \le \mu_{(k+1)M+r_0} f_{l((k+1)M+r_0)} - \sigma_{(k+1)M+r_0}$$

From

$$kM + r_0 + 1 \le l((k+1)M + r_0) \le (k+1)M + r_0$$

$$f_{l((k+1)M+r_0)} \le |f_0| e^{\lambda s(l((k+1)M+r_0)-1)} - e^{-\lambda s(l(k+1)M+r_0)-1)} \sum_{i=0}^{(k+1)} \sigma_{t(i)}$$

$$f_{(k+1)M++r_0+1} \le \mu_{(k+1)M+r_0} \{ |f_0| e^{\lambda s(l((k+1)M+r_0)-1)} - e^{-\lambda s(l(k+1)M+r_0)-1)} \sum_{i=0}^{k+1} \sigma_{t(i)} \} - \sigma_{(k+1)M+r_0}$$

$$\le e^{\lambda h_{(k+1)M+r_0}} |f_0| e^{\lambda s(((k+1)M+r_0)-1)} - e^{-\lambda h_{(k+1)M+r_0}} e^{-\lambda s((k+1)M+r_0)-1)} \sum_{i=0}^{k+1} \sigma_{t(i)} - \mu_{(k+1)M+r_0} e^{-\lambda s((k+1)M+r_0)-1)} \sigma_{(k+1)M+r_0}$$

$$\le e^{\lambda h_{(k+1)M+r_0}} |f_0| e^{\lambda s(((k+1)M+r_0)-1)} - e^{-\lambda h_{(k+1)M+r_0}} e^{-\lambda s((k+1)M+r_0)-1)} \sum_{i=0}^{k+1} \sigma_{t(i)} - e^{-\lambda h_{(k+1)M+r_0}} e^{-\lambda s((k+1)M+r_0)-1)} \sigma_{(k+1)M+r_0}$$

$$f_{(k+1)M+r_0+1} \le |f_0| e^{\lambda s((k+1)M+r_0)} - e^{-\lambda s((k+1)M+r_0))} \sum_{i=0}^{k+1} \sigma_{t(i)}$$

Therefore (*) holds for $r=r_0+1$.

This implies (*) is true for every r and every k.

Thus

$$f_{jM+r} \le |f_0| e^{\lambda s(jM+r-1)} - e^{-\lambda s(jM+r-1)} \sum_{i=0}^{j} \sigma_{t(i)}$$

T1: *Under A1 and A2, let the search direction* d_k *and the step-length* α_k *be determined by A2 and (1), respectively. Then,* $\lim_{k\to\infty} \inf \|g_k\| = 0$.

Proof: Recalling $\sum_{i=0}^{\infty} h_i = \gamma$, we derive from L3 that

$$f_{jM+r} \le |f_0| e^{\lambda r} - e^{-\lambda r} \sum_{i=0}^{j} \sigma_{t(i)}$$

with $t(i) \in \{iM + r - 1 | 1 \le r \le M\}$. As a consequence, we have

$$0 \le e^{-\lambda r} \sum_{i=0}^{j} \sigma_{t(i)} \le e^{\lambda r} |f_0| - f_{jM+r} \qquad (4)$$

By A1, f_k is bounded on the level set $\pounds_c$. Combined with the fact that $e^{\pm\lambda r}$ is nonzero and finite, (4) indicates that $\sigma_{t(i)}$ is also finite. It follows that

$$\lim_{i\to\infty} \sigma_{t(i)} = 0,$$

When $j \to \infty$. In view of non-monotone F-rule and the definition of the forcing function, we have

$$\lim_{i\to\infty} t(i) = \lim_{i\to\infty} \left(-\frac{g_i^T d_i}{||d_i||} \right) = 0.$$

That is,

$$\lim_{i\to\infty} ||g_i^T||\cos(-g_i^T, d_i)=0.$$

By A2,

$$\lim_{i\to\infty} \cos(-g_i^T, d_i) \neq 0,$$

and therefore, we obtain

$$\lim_{i\to\infty} \left||g_i^T|\right| = 0 \text{ (or) } \lim_{k\to\infty} ||g_k|| = 0.$$

The proof is completed.

2.1.5 Order of Convergence

We know that $x_{k+1} = x_k + \alpha_k d_k$

Let ξ be the actual root of the function

$$\begin{aligned} |\xi - x_{k+1}| &= |\xi - x_k - \alpha_k d_k| \\ &\leq |\xi - x_k| + |\alpha_k d_k| \\ |\xi - x_{k+1}| &\leq |\xi - x_k| \end{aligned}$$

Thus the order of convergence of this method is 1 and the rate of linear convergence is also 1.

3.Numerical Tests

3.1 Algorithm:

Step 1: Given the initial values $x_0 \in R^n$ and the other data including an integer

M≥ 1, a constant

$\varepsilon > 0$, $\rho\epsilon(0,0.5)$,a s.p.d matrix $H_0 \epsilon R^{n\times n}$, $\psi\epsilon(0,1)$, as well as $k=0$.

Step 2: Examine the stopping criterion by computing $g_k = \nabla f(x_k)$. If $\left||g_k|\right| \leq \varepsilon$,

$x^*=x_k$ and the algorithm stops

Step 3: Set $\alpha = 1$.Compute the search direction

$$d_k = \begin{cases} -g_k & if\ g_k^T d_k < \psi ||g_k||^2 \\ -H_k g_k & otherwise \end{cases}$$

Step 4: Let *m(k)={k,M-1}* and μ_k satisfy (1).

If

$$f_{k+1} \le \mu_k \left\{ \max_{0\le r\le m(k)} f(x_{k-r}) \right\} + \rho\alpha g_k^T d_k$$

holds, $\alpha_k = \alpha$. Otherwise contract α using quadratic interpolation method.

Step 5: Compute the next point. Set $s_k = \alpha_k d_k$ and $x_{k+1} = x_k + s_k$. Then, compute

$f(x_{k+1})$ and $g_{k+1} = \nabla f(x_{k+1})$.

Step 6: Update the iteration matrix H_k, using BFGS formula. Set *k=k+1*and then go

to the next step.

Take the parameters involved in the algorithm as

$\varepsilon = 10^{-3}, \rho = 10^{-3}, H_0 = I, \psi = 10^{-5}, \mu_k = e^{\lambda h_k}$, M=1 and λ=1 with $h_k = \frac{sign\ \{f_{l(k)}\}}{(k+1)^2}$ where$\lambda \ge 1$.The initial values are taken in such a way that they are nearer to the actual values.

Table 1:

Function	**Initial Values**	**Improved Non-Monotone Line Search Slackness Technique**	**Non-Monotone Line Search Slackness Technique**	**Actual Value**
$x^3+3xy^2-15x^2-15y^2+72x$	*x=4.5* *y=0.5*	***x=3.9998779147*** ***y=-0.00022085215480*** ***Max=111.9999999105***	***x=4.000665210*** ***y=-0.00009131169777*** ***Max=111.9999986477***	***x=4*** ***y=0*** ***Max=112***
		No. of Iterations=7	**No. of Iterations=43**	
$x^3+y^3-12x-3y+20$	*x=-2* *y=-0.95*	***x=-2*** ***y=-0.9994004*** ***Max=37.99999892***	***x=-2*** ***y=-1.00081432711*** ***Max=37.999998010074***	***X=-2*** ***Y=-1*** ***Max=38***
		No. of Iterations=3	**No. of Iterations=4**	
$2(x-y)^2-x^4-y^4$	*x=1.5*	***x=1.414217719***	***x=1.414217719***	***x=***

	y=-1.5	*y=-1.414217719* *Max=7.9999999972355* **No. of Iterations=6**	*y=-1.414217719* *Max=7.9999999972355* **No. of Iterations=6**	*1.414213* *y=-1.414213* *Max=8*
$x^3y^2(1-x-y)$	*x=0.6* *y=0.3*	*x=0.59028* *y=0.29352* *Max=0.002058997* **No. of Iterations=1**	*x=0.59028* *y=0.29352* *Max=0.002058997* **No. of Iterations=1**	*x=0.5* *y=0.3333* *Max=0.0023*
xy(3x+2y+1)	*x=-0.1* *y=-0.15*	*X=-0.115* *Y=-0.16* *Max=0.006164* **No. of Iterations=1**	*X=-0.115* *Y=-0.16* *Max=0.006164* **No. of Iterations=1**	*X=-0.11111* *Y=-0.16667* *Max=0.0062*

Conclusion

In this paper, any unconstrained function is considered and the accurate maximum value is found by using an improved non monotone line search slackness technique, which possess convergence. The effectiveness of the algorithm is proved by the numerical results.

References

1. Ping Hu,Xu-Qing Liu: A Nonmonotone Line search slackness technique for unconstrained optimization.J Optim Theory Appl(2013),158:773-786
2. Engineering optimization
3. Grippo, L., Lampariello, F., Lucidi, S.: A nonmonotone line search technique for Newton’s methods. SIAM J. Numer. Anal. **23,** 707–716 (1986)
4. Sun, W., Han, J., Sun, J.: Global convergence of non-monotone descent methods for unconstrained optimization problems. J. Comput. Appl. Math. **146,** 89–98 (2002)
5. Yu, Z., Pu, D.: A new nonmonotone line search technique for unconstrained optimization. J. Comput. Appl. Math. **219,** 134–144 (2008)
6. Grippo, L., Lampariello, F., Lucidi, S.: A truncated Newton method with nonmonotone line search for unconstrained optimization. J. Optim. Theory Appl. **60**,401–419 (1989)
7. Nocedal, J., Wright, S.J.: Numerical Optimization, 2nd edn. Springer, New York (2006)
8. Shi, Z.J., Shen, J.: New inexact line search method for unconstrained optimization. J. Optim. Theory Appl. **127,** 425–446 (2005)